

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-20

Perfect score: 60

Sequence: 1 PQGVLRVRFV 12

Scoring table: BLOSUM62, Gapex 0.5
Gapop 10.0, Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_8.*

1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*
10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	12	4 AAB35379	Aab35379 Alpha3bet
2	55	91.7	12	4 AAB35352	Aab35352 Alpha3bet
3	55	91.7	12	4 AAB35378	Aab35378 Alpha3bet
4	55	91.7	12	4 AAB35376	Aab35376 Alpha3bet
5	55	91.7	12	4 AAB35380	Aab35380 Alpha3bet
6	55	91.7	12	6 AAB35380	Abg72834 Thrombos
7	55	91.7	240	8 ADL70641	Adl70641 Human thr
8	55	91.7	432	8 ADQ39359	Adg39359 Human myo
9	55	91.7	432	8 ADQ39357	Adg39357 Human myo
10	55	91.7	459	4 AAU02316	Aau02316 Angiotens
11	55	91.7	466	3 AAB43602	Aab43602 Human can
12	55	91.7	546	4 AAU02915	Aau02915 Angiotens
13	55	91.7	548	7 AAU02474	Adn02474 TSF polyp
14	55	91.7	555	4 AAU02914	Aau02914 Angiotens
15	55	91.7	731	4 AAU02913	Aau02913 Angiotens
16	55	91.7	1152	3 AAB00042	Aab00042 Human thr
17	55	91.7	1152	5 AAU74771	Aau74771 Human thr
18	55	91.7	1152	5 ABB82285	Abb82285 Human thr
19	55	91.7	1170	4 AAB74450	Aab74450 Human var
20	55	91.7	1170	4 AAB90800	Aab90800 Human she
21	55	91.7	1170	5 AAE25030	Aae25030 Human thr
22	55	91.7	1170	5 AAU75315	Aau75315 Human thr
23	55	91.7	1170	6 ABP96780	Abp96780 Human COP

ALIGNMENTS

RESULT 1

AAB35379
ID AAB35379 standard; peptide; 12 AA.
XX
AC AAB35379;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3bet integrin binding peptide #44.
XX
KW Alpha3bet integrin; angiogenesis; cell proliferation; cancer;
diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
useful for inhibiting cell adhesion to extracellular matrix, cell
motility and proliferation and for treating rheumatoid arthritis and
cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
alpha3beta1 integrins. They are useful in the modulation of cell adhesion
and motility, and in the treatment of cancer, diabetic retinopathy,
rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
and restenosis. The present sequence is an example of one of the peptides
of the invention
XX
SQ Sequence 12 AA;

24	55	91.7	1170	6	ABU03474	Abu03474 Angiogene
25	55	91.7	1170	6	ABG74673	Abg74673 Human THB
26	55	91.7	1170	6	AAE36228	AAe36228 Human THB
27	55	91.7	1170	7	ABR62059	ABr62059 Human thr
28	55	91.7	1170	7	ADN39852	ADn39852 Cancer/an
29	55	91.7	1170	8	ADJ76124	Adj76124 Marker ge
30	55	91.7	1170	8	ADJ75296	Adj75296 Marker ge
31	55	91.7	1170	8	ADL70639	Adl70639 Human thr
32	55	91.7	1170	8	ADL35874	Adl35874 Human thr
33	55	91.7	1170	8	ADQ26070	Adq26070 Thrombos
34	55	91.7	1170	8	ADP54179	Adp54179 Human PRO
35	55	91.7	1170	8	ADQ39358	Adg39358 Human myo
36	55	91.7	1170	8	ADQ39356	Adg39356 Human myo
37	55	91.7	1170	8	ADQ39355	Adg39355 Human myo
38	55	91.7	1170	8	ADZ21688	Adz21688 Thrombos
39	55	91.7	1170	9	ABE87781	Aeb87781 Human thr
40	55	91.7	1170	9	ABE46751	Aeb46751 Human thr
41	54	90.0	12	4	AAB35361	Aab35361 Alpha3bet
42	52	86.7	12	4	AAB35373	Aab35373 Alpha3bet
43	52	86.7	12	4	AAB35381	Aab35381 Alpha3bet
44	51	85.0	12	4	AAB35364	Aab35364 Alpha3bet
45	51	85.0	12	4	AAB35374	Aab35374 Alpha3bet

Query Match 100.0%; Score 60; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQLQVRFFV 12
Db 1 FQGVQLQVRFFV 12

RESULT 2
AAB35352
ID AAB35352 standard; peptide; 12 AA.

XX
AC AAB35352;
XX
DT 08-MAY-2001 (first entry)
XX

DE Alpha3betal integrin binding peptide #17.

KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX
OS Synthetic.

XX WO200105812-A2.

XX PN
XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX DR WPI; 2001-182656/18.

XX PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.

XX PS Claim 4; Page 34; 84pp; English.

XX CC The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention

XX SQ Sequence 12 AA;

Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQLQVRFFV 12
Db 1 FQGVQLQVRFFV 12

RESULT 3
AAB35378
ID AAB35378 standard; peptide; 12 AA.

XX
AC AAB35378;

XX DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #43.

XX Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX DR WPI; 2001-182656/18.

XX PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.

XX PS Example 2; Page 34; 84pp; English.

XX CC The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention

XX SQ Sequence 12 AA;

Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQLQVRFFV 12
Db 1 FQGVQLQVRFFV 12

RESULT 4
AAB35376

ID AAB35376 standard; peptide; 12 AA.

XX AC AAB35376;

XX DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #41.

XX KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX WPI; 2001-182656/18.
XX New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX Claim 4; Page 34; 84pp; English.
XX The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;

Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVVF 12
| | | | | | | | | |
Db 1 FQGVLOQVRVVF 12

RESULT 5
AAB35380
ID AAB35380 standard; peptide; 12 AA.
AC AAB35380;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3betal integrin binding peptide #45.
XX
KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX Synthetic.
OS
XX WO200105812-A2.
XX
XX 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
XX 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Krutzsch HC;
XX
XX WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
XX Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;

Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVVF 12
| | | | | | | | | |
Db 1 FQGVLOQVRVVF 12

RESULT 6
ABG72834
ID ABG72834 standard; peptide; 12 AA.
XX
AC ABG72834;
XX
DT 24-FEB-2003 (first entry)
XX
DE Thrombospondin-1 sequence containing synthetic peptide.
XX
KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX Synthetic.
OS
XX WO2002081630-A2.
XX
XX 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010535.
XX
XX 06-APR-2001; 2001US-0281994P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Krutzsch HC;
XX
XX WPI; 2003-103329/09.
XX
PT A new diagnosis for cancer other than prostate cancer in a mammal useful
PT to detect cancer including lung cancer, particularly small cell lung
PT cancer and melanoma comprises detecting semenogelin in a sample.
XX
XX Example 1; Page 14; 32pp; English.
XX
CC The invention relates to diagnosing cancer other than prostate cancer in
CC a male mammal, comprising assaying a test sample for increased level of
CC semenogelin, or cancer in a female by assaying for the presence of
CC semenogelin. Administering a semenogelin protein or polypeptide fragment
CC or a semenogelin-specific antibody or active fragment, or a recombinant
CC vector expressing the protein or antibody, is useful for inducing an
CC immune response to a cancer in a mammal, where the cancer is not prostate
CC cancer and semenogelin is a marker. The invention is used to diagnose
CC cancer, particularly of epithelial origin such as lung cancer, papillary
CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
CC (SCLC), or a melanoma. The present sequence represents the amino acid
CC sequence of the thrombospondin-1 sequence containing synthetic peptide
CC which binds to alpha-3-beta-1 integrin
XX
SQ Sequence 12 AA;

Query Match 91.7%; Score 55; DB 6; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVVF 12
| | | | | | | | | |
Db 1 FQGVLOQVRVVF 12

RESULT 7
ADL70641

ID ADL70641 standard; protein; 240 AA.
 XX ADL70641;
 XX
 XX
 DT 20-MAY-2004 (first entry)
 XX
 XX Human thrombospondin-1 N-terminal domain.
 DE
 XX Human; thrombospondin-1; epitope; cancer; diagnosis.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 23..32 /note= "Heparin binding region"
 FT Region 77..82 /note= "Heparin binding region"
 FT Region 151..164 /note= "Fibrinogen binding region"
 FT
 XX WO2004018995-A2.
 XX
 XX 04-MAR-2004.
 XX
 XX 20-AUG-2003; 2003WO-US026023.
 XX
 XX 23-AUG-2002; 2002US-0405494P.
 PR 21-APR-2003; 2003US-00419462.
 XX
 XX (WILL/) WILLIAMS K J.
 PA
 XX Williams KJ;
 PI
 XX WPI; 2004-226901/21.
 DR
 XX
 XX New purified thrombospondin fragment extracted from a body fluid, useful
 FT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
 FT or leukemia or as calibrators, indicators, immunogens and analytes.
 FT
 XX Disclosure; SEQ ID NO 40; 76pp; English.
 PS
 XX
 XX The present sequence is that of the N-terminal domain of human
 CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
 CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
 CC clinical assays for cancer and for generation of antibodies and other
 CC binding agents. A method that distinguishes TSP from a TSP fragment or
 CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
 CC or portion as a target for a binding molecule, e.g. an antibody, to
 CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
 CC epitope present in TSP but not in the fragment or portion to obtain a
 CC quantitation of TSP only; and (3) using the difference between (1) and
 CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
 CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
 CC the TSP fragment or portion is performed in order to detect the presence,
 CC or monitor the course, of a disease or condition selected from cancer,
 CC renal failure, renal disease, atopic dermatitis, vasculitis, acute
 CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
 CC infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
 CC nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
 CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
 CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
 CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition
 CC associated with clotting, a condition associated with platelet
 CC activation, a condition associated with intravascular platelet
 CC activation, a condition associated with consumption of platelets, heparin
 CC -induced thrombocytopenia, disseminated intravascular coagulation,
 CC intravascular coagulation, extravascular coagulation, a condition
 CC associated with endothelial activation, a condition associated with
 CC production and/or release of thrombospondin and/or a thrombospondin
 CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 CC reaction, an aspartame reaction, atopic dermatitis, eczema,
 CC hypersensitivity, scleroderma, conditions associated with plugging of
 CC vessels, a condition associated with a cryofibrinogen, a condition

CC associated with a cryoglobulin, and a condition associated with an anti-
 CC cardioplin antibody. The cancer is selected from adenoma,
 CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system,
 CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
 CC biliary system, pancreas, head, neck, endocrine system, reproductive
 CC system (male or female), genitourinary system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.
 XX
 SQ Sequence 240 AA;
 Query Match 91.7%; Score 55; DB 8; Length 240;
 Best Local Similarity 91.7%; Pred. No. 0.033;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVLQVRVFV 12
 Db 190 FQGVLQVRVFV 201
 RESULT 8
 ADQ39359
 ID ADQ39359 standard; protein; 432 AA.
 XX
 AC ADQ39359;
 XX
 XX 18-NOV-2004 (first entry)
 DT
 XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.
 DE
 XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 KW cardiatic; gene therapy; human.
 KW
 XX Homo sapiens.
 OS
 XX WO2004058052-A2.
 PN
 XX 15-JUL-2004.
 PD
 XX
 XX 22-DEC-2003; 2003WO-US040978.
 PF
 XX 20-DEC-2002; 2002US-0434778P.
 PR 10-MAR-2003; 2003US-0453135P.
 PR 30-APR-2003; 2003US-0466412P.
 PR 23-SEP-2003; 2003US-0504955P.
 PR
 XX (APPL-) APPLERA CORP.
 PA
 XX Cargill M, Devlin JJ, Iakoubova O;
 PI
 XX WPI; 2004-533949/51.
 DR N-PSDB; ADQ38531.
 DR
 XX Identifying an individual who has an altered risk for developing
 FT myocardial infarction by detecting a single nucleotide polymorphism in
 FT the individual's nucleic acids.
 FT
 XX Claim 10; SEQ ID NO 1022; 145pp; English.
 PS
 XX The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino

CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC sequence represents the protein of a human myocardial infarction.
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.

XX
 SQ Sequence 432 AA;

Query Match 91.7%; Score 55; DB 8; Length 432;

Best Local Similarity 91.7%; Pred. No. 0.063;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLOQVRVF 12

DB 208 FQGVLOQVRVF 219

RESULT 9

ADQ39357

ID ADQ39357 standard; protein; 432 AA.

XX

AC ADQ39357;

XX

DT 18-NOV-2004 (first entry)

XX

DE Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

XX

KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW

KW cardiant; gene therapy; human.

XX

OS Homo sapiens.

XX

PN WO2004058052-A2.

XX

PD 15-JUL-2004.

XX

PF 22-DEC-2003; 2003WO-US040978.

XX

PR 20-DEC-2002; 2002US-0434778P.

XX

PR 10-MAR-2003; 2003US-0453135P.

XX

PR 30-APR-2003; 2003US-0466412P.

XX

PR 23-SEP-2003; 2003US-0504955P.

XX

PA (APPL-) APPLERA CORP.

XX

PI Cargill M, Devlin J, Iakoubova O;

XX

XX WPI; 2004-533949/51.

XX

DR N-PSDB; ADQ38529.

XX

PT Identifying an individual who has an altered risk for developing

PT

PT myocardial infarction by detecting a single nucleotide polymorphism in

PT

PT the individual's nucleic acids.

XX

PS Claim 10; SEQ ID NO 1020; 145pp; English.

XX

XX The invention relates to a novel method for identifying an individual who

CC

CC has an altered risk for developing myocardial infarction. The method

CC

CC comprises detecting a single nucleotide polymorphism (SNP) in any one of

CC

CC the nucleotide sequences given in the specification in the individual's

CC

CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC sequence represents the protein of a human myocardial infarction.
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.

XX
 SQ Sequence 432 AA;

Query Match 91.7%; Score 55; DB 8; Length 432;

Best Local Similarity 91.7%; Pred. No. 0.063;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLOQVRVF 12

DB 208 FQGVLOQVRVF 219

RESULT 10

AAU02916

ID AAU02916 standard; protein; 459 AA.

XX

AC AAU02916;

XX

DT 12-SEP-2001 (first entry)

XX

DE Angiotensin converting enzyme (ACEV) splice variant protein #16.

XX

KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidetic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.

XX

OS Homo sapiens.

XX

PN WO200136632-A2.

XX

PD 25-MAY-2001.

XX

PF 17-NOV-2000; 2000WO-IL000766.

XX

PR 17-NOV-1999; 99IL-00132978.

XX

PR 10-DEC-1999; 99IL-00133455.

XX

PA (COMP-) COMPUGEN LTD.

XX

PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX

XX WPI; 2001-336004/35.

DR

DR N-PSDB; AAS06016.

XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 16; 519pp; English.
XX
CC The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
XX Sequence 459 AA;
XX
XX Query Match 91.7%; Score 55; DB 4; Length 459;
XX Best Local Similarity 91.7%; Pred. No. 0.067;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLLQVRFVF 12
XX ||||| |||||
XX Db 208 FQGVLLQVRFVF 219
XX
XX RESULT 11
XX AAB43602
XX ID AAB43602 standard; protein; 466 AA.
XX AC AAB43602;
XX XX
XX DT 08-FEB-2001 (first entry)
XX XX
XX DE Human cancer associated protein sequence SEQ ID NO:1047.
XX XX
XX KW Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
XX antidiabetic; antiaesthatic; antirheumatic; antithratic; antiviral;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening.
XX
XX OS Homo sapiens.
XX XX
XX PN WO20005350-A1.
XX XX
XX PD 21-SEP-2000.
XX XX
XX PF 08-MAR-2000; 2000WO-US005882.
XX XX
XX PR 12-MAR-1999; 99US-0124270P.
XX XX
XX XX (HUMA-) HUMAN GENOME SCI INC.
XX PA Rosen CA, Ruben SM;
XX PI
XX DR WPI; 2000-587533/55.
XX DR N-PSDB; AAC77811.
XX XX

PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
PS Claim 11; Page 1636-1638; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiaesthatic; antirheumatic; antithratic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
XX Sequence 466 AA;
XX
XX Query Match 91.7%; Score 55; DB 3; Length 466;
XX Best Local Similarity 91.7%; Pred. No. 0.068;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLLQVRFVF 12
XX ||||| |||||
XX Db 261 FQGVLLQVRFVF 272
XX
XX RESULT 12
XX AAU02915
XX ID AAU02915 standard; protein; 546 AA.
XX AC AAU02915;
XX XX
XX DT 12-SEP-2001 (first entry)
XX XX
XX DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX XX
XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX platelet-derived endothelial cell growth factor; cardiovascular disease;
XX cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX myocardial infarction; coronary arterial thrombosis; renal disease;
XX diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
XX multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
XX vascular disorder; asbestosis.
XX
XX OS Homo sapiens.
XX XX
XX PN WO200136632-A2.
XX XX
XX PD 25-MAY-2001.
XX XX
XX PF 17-NOV-2000; 2000WO-IL000766.
XX XX
XX PR 17-NOV-1999; 99IL-00132978.
XX PR 10-DEC-1999; 99IL-00133455.
XX XX
XX PA (COMP-) COMPUGEN LTD.
XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.
DR N-PSDB; AAS06015.
XX
PT Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 15; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 546 AA;
Query Match 91.7%; Score 55; DB 4; Length 546;
Best Local Similarity 91.7%; Pred. No. 0.08;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLOQVRFFV 12
Db 208 FQGVLOQVRFFV 219
|||||
RESULT 13
ADN02474
ID ADN02474 standard; protein; 548 AA.
XX
AC ADN02474;
XX
XX 17-JUN-2004 (first entry)
DT TSF polypeptide.
XX
DE adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX
KW Homo sapiens.
XX
OS CN1401387-A.
XX
PN 12-MAR-2003.
XX
PD 21-AUG-2002; 2002CN-00129408.
XX
PF 21-AUG-2002; 2002CN-00129408.
XX
PR (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
XX
PA Han Z, Liu P;
XX
PI WPI; 2003-469302/45.
XX
DR N-PSDB; ADN02475.
XX
PT Tumor suppressing polypeptide TSF and gene therapy vector composition.
XX
PS Claim 2; SEQ ID NO 1; 13pp; Chinese.
XX
XX The present invention relates to a novel recombinant adenovirus vector
CC mediated anti-neoplastic composition is prepared through cloning the CDNA

CC sequence from the human peripheral blood cell by specific primer and
CC reverse transcription-polymerase chain reaction (RT-PCR) method for
CC coding TSF polypeptide, construction in human embryonic kidney cell 293
CC by AdEasy system, and packaging and expressing the recombinant adenovirus
CC vector of TSF. It can suppress the growth and transfer of cancer. The
CC present sequence represents the TSF polypeptide.
XX
SQ Sequence 548 AA;
Query Match 91.7%; Score 55; DB 7; Length 548;
Best Local Similarity 91.7%; Pred. No. 0.081;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLOQVRFFV 12
Db 208 FQGVLOQVRFFV 219
|||||
RESULT 14
AAU02914
ID AAU02914 standard; protein; 555 AA.
XX
AC AAU02914;
XX
XX 12-SEP-2001 (first entry)
DT Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
DE Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX
OS Homo sapiens.
XX
PN WO200136632-A2.
XX
XX 25-MAY-2001.
PD 17-NOV-2000; 2000WO-IL000766.
XX
PF 17-NOV-1999; 99IL-00132978.
XX
PR 10-DEC-1999; 99IL-00133455.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX
XX WPI; 2001-336004/35.
XX
DR N-PSDB; AAS06014.
XX
XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
XX Claim 4; Fig 14; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,

CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 555 AA;

Query Match 91.7%; Score 55; DB 4; Length 555;
 Best Local Similarity 91.7%; Pred. No. 0.082;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVVF 12
 ||||| |||||
 Db 208 FQGVLOQVRVVF 219

RESULT 15

AAU02913
 ID AAU02913 standard; protein; 731 AA.

AC AAU02913;

XX 12-SEP-2001 (first entry)

XX Angiotensin converting enzyme (ACEV) splice variant protein #13.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.

XX Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

XX 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

XX N-PSDB; AAS06013.

XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 4; Fig 13; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,

CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 731 AA;

Query Match 91.7%; Score 55; DB 4; Length 731;
 Best Local Similarity 91.7%; Pred. No. 0.11;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVVF 12
 ||||| |||||
 Db 208 FQGVLOQVRVVF 219

Search completed: June 5, 2006, 22:24:58

Job time : 91.1379 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-20

Perfect score: 60

Sequence: 1 FQGVLQVRVVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	91.7	229	2	thrombospondin 1 -
2	55	91.7	1170	1	thrombospondin 1 p
3	55	91.7	1170	2	thrombospondin 1 p
4	38	63.3	212	2	probable imidazole
5	38	63.3	469	2	hypothetical prote
6	37	61.7	151	2	tRNA-pseudouridine
7	37	61.7	1172	1	thrombospondin 2 p
8	37	61.7	1172	2	thrombospondin 2 p
9	36	60.0	102	2	hypothetical prote
10	36	60.0	145	2	hypothetical prote
11	36	60.0	162	2	protein F15D4.3 [i
12	36	60.0	1299	2	two-component hybr
13	35	58.3	102	2	hypothetical prote
14	35	58.3	313	2	NSP3 protein - hum
15	35	58.3	471	2	53K glycoprotein -
16	35	58.3	516	2	methylnalonyl-CoA
17	35	58.3	527	2	hypothetical prote
18	35	58.3	893	2	probable penicilli
19	35	58.3	1142	1	M polypeptide prec
20	35	58.3	1148	1	M polypeptide prec
21	35	58.3	1178	1	xanthine dehydroge
22	35	58.3	1333	1	hypothetical prote
23	34	56.7	165	2	hypothetical prote
24	34	56.7	175	2	35K GTP-binding pr
25	34	56.7	295	2	ribose/galactose A
26	34	56.7	311	2	hypothetical prote
27	34	56.7	400	2	hypothetical prote
28	34	56.7	402	2	amino acid permea
29	34	56.7	459	2	

RESULT 1

S57957
thrombospondin 1 - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S57957
R:Lafuente, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.
submitted to the EMBL Data Library, July 1995
A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi
A:Reference number: S57955
A:Accession: S57957
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <LAF>
A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:g899228; PIDN
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; voi

Query Match 91.7%; Score 55; DB 2; Length 229;
Best Local Similarity 91.7%; Pred. No. 0.0017;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLQVRVVF 12
Db 190 FQGVLQVRVVF 201
|||||

RESULT 2

TSHUP1
thrombospondin 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: A26155; A34274; A30140; A25812; A05172; A42927
R:Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c
A:Reference number: A26155; MUID:87057617; PMID:2430973
A:Accession: A26155
A:Molecule type: mRNA
A:Residues: 1-1170 <LAW>
A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:g37137; PIDN:CI
A:Note: parts of this sequence, including the amino end of the mature protein, were deter
R:Lahterty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA se
A:Reference number: A34274; MUID:89291870; PMID:2544587
A:Accession: A34274
A:Molecule type: DNA
A:Residues: 1-166 <LAH>
A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835
R:Hennessey, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I
J. Cell Biol. 108, 729-736, 1989
A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

ALIGNMENTS

30 34 56.7 477 1 P2WLEP L2 protein - Europ
31 34 56.7 505 2 E87021 probable integral-
32 34 56.7 552 2 A70709 probable ptrBa pro
33 34 56.7 565 2 T23843 hypothetical prote
34 34 56.7 683 2 AC0741 oligopeptidase B (
35 34 56.7 686 1 E64946 oligopeptidase B (
36 34 56.7 686 2 C90948 proteinase II (imp
37 34 56.7 686 2 G85796 proteinase II (imp
38 34 56.7 724 2 F87187 proteinase II (imp
39 34 56.7 747 2 AE2929 two component resp
40 34 56.7 783 2 A98353 probable transcrip
41 34 56.7 913 2 T52485 neurofilament prot
42 34 56.7 1142 2 T39103 probable negative
43 34 56.7 1312 2 S68593 DNA-directed DNA p
44 34 56.7 1670 2 S71551 DNA-directed DNA p
45 33 55.0 82 2 F96625 hypothetical prote

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: hsh; NWA0840

C;Superfamily: amidotransferase hsh; trpG homology

C;Keywords: glycosyltransferase; pentosyltransferase

Query Match 63.3%; Score 38; DB 2; Length 212;

Best Local Similarity 58.3%; Pred. No. 4.6;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRPFF 12

|||:|:|:|

Db 144 FQGIQDTRPFF 155

RESULT 5

T33595

hypothetical protein C40A11.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T33595

R;Maggi, L.; Goela, D.

submitted to the EMBL Data Library, October 1998

A;Description: The sequence of C. elegans cosmid C40A11.

A;Reference number: Z21374

A;Accession: T33595

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-469 <MAG>

A;Cross-references: UNIPROT:Q9TZA3; UNIPARC:UPI000007BD5F; EMBL:AF099914; PIDN:AAC68764.

A;Experimental source: strain Bristol N2; clone C40A11

C;Genetics:

A;Gene: CESP:C40A11.9

A;Map position: 2

A;Introns: 199/2; 315/2; 424/2.

Query Match

Best Local Similarity 63.3%; Score 38; DB 2; Length 469;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRPFF 12

|||:|:|:|

Db 347 FNGVVTQIYFVF 358

RESULT 6

C57253

tRNA-pseudouridine synthase (EC 5.4.99.-) - Acinetobacter calcoaceticus (fragment)

N;Alternate names: hypothetical protein lipB 5'-region

C;Species: Acinetobacter calcoaceticus

C;Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 05-Oct-2004

C;Accession: C57253

R;Kok, R.G.; van Thor, J.J.; Nugteren-Roodzant, I.M.; Vosman, B.; Hellingwerf, K.J.

J. Bacteriol. 177, 3295-3307, 1995

A;Title: Characterization of lipase-deficient mutants of Acinetobacter calcoaceticus BD4

A;Reference number: A57253; MUID:95286514; PMID:7768830

A;Accession: C57253

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-151 <KOK>

A;Cross-references: UNIPARC:UPI000017879A; GB:X80800

C;Superfamily: tRNA pseudouridine synthase B

C;Keywords: intramolecular transferase; isomerase; tRNA modification

Query Match

Best Local Similarity 61.7%; Score 37; DB 2; Length 151;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GVLQQVRPFF 12

|||:|:|:|

Db 29 GVLQKVRWLF 38

RESULT 7

TSHP2

thrombospondin 2 precursor - human

C;Species: Homo sapiens (man)

C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C;Accession: A47379; A42173

R;LaBell, T.L.; Byers, P.H.

Genomics 17, 225-229, 1993

A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: poter

A;Reference number: A47379; MUID:94010892; PMID:8406456

A;Accession: A47379

A;Molecule type: mRNA

A;Residues: 1-1172 <LAB>

A;Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046580; GB:L12350; NID:G307505; PIDN:1

R;LaBell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.

Genomics 12, 421-429, 1992

A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression of

A;Reference number: A42173; MUID:92217961; PMID:1559694

A;Accession: A42173

A;Molecule type: mRNA

A;Residues: 560-1172 <LA2>

A;Cross-references: UNIPARC:UPI00001742C1; GB:M81339

A;Experimental source: fibroblast

A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)

C;Genetics:

A;Gene: GDB:THBS2; TSP2

A;Cross-references: GDB:128789; OMIM:188061

A;Map position: 6q27-6q27

C;Complex: homotrimer, disulfide linked

C;Function:

A;Description: participates in cell migration and adhesion, and in platelet aggregation

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1172/Product: thrombospondin 2 #status predicted <MAT>

F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>

F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>

F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>

F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>

F;553-588/Domain: EGF homology <EGF1>

F;652-691/Domain: EGF homology <EGF>

F;928-930/Region: cell attachment (R-G-D) motif

F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predic

F;167-226/Disulfide bonds: #status predicted

F;266,270/Disulfide bonds: interchain #status predicted

F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match

Best Local Similarity 61.7%; Score 37; DB 1; Length 1172;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRPFF 12

|||:|:|:|

Db 202 FRGLQNVHLVF 213

RESULT 8

A42587

thrombospondin 2 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A42587; A39851

R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.

J. Biol. Chem. 267, 3274-3281, 1992

A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell

A;Reference number: A42587; MUID:92147683; PMID:1371115

A;Accession: A42587

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-1172 <LAH>

A;Cross-references: UNIPROT:Q03350; UNIPARC:UPI0000029847; GB:L07803; GB:M87275; NID:G34

A;Note: sequence extracted from NCBI backbone (NCBIP:81502)

R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.

J. Biol. Chem. 266, 12821-12824, 1991
A;Title: A second, expressed thrombospondin gene (Thb2) exists in the mouse genome.
A;Reference number: A39851; MUID:91302287; PMID:1712771
A;Accession: A39851
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-873 <BOR>
A;Cross-references: UNIPARC:UPI000016D077; GB:M64866; NID:q201994; PIDN:AAA0432.1; PID:
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
F;319-377/Domain: von Willebrand factor type C repeat homology <VMC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>

Query Match 61.7%; Score 37; DB 2; Length 1172;
Best Local Similarity 59.3%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLQQVRVFV 12
|:|:|:|:|
Db 202 FRGLQNVHLVF 213

RESULT 9
D81910
hypothetical protein NMA1411 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jäglele, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:2022556; PMID:10761919
A;Accession: D81910
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-102 <PAR>
A;Cross-references: UNIPROT:Q9JUB1; UNIPARC:UPI00000C4BB6; GB:AL162755; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1411
C;Superfamily: Neisseria meningitidis hypothetical protein NMB1242

Query Match 60.0%; Score 36; DB 2; Length 102;
Best Local Similarity 77.8%; Pred. No. 5.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLLQQVRVF 10
|:|:|:|:|
Db 54 QAVVQQVRVF 62

RESULT 10
T20985
hypothetical protein F15D4.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20985
R;Smyle, R.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19354
A;Accession: T20985
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-145 <WIL>
A;Cross-references: UNIPROT:Q93511; UNIPARC:UPI000007BE33; EMBL:Z80344; PIDN:CAB02486.2;
A;Experimental source: clone F15D4
C;Genetics:
A;Gene: CESP:F15D4.3

A;Map position: 2
A;Introns: 21/3; 82/1

Query Match 60.0%; Score 36; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLLQQVRVFV 12
|:|:|:|:|
Db 73 FMGVAQGLRYIF 84

RESULT 11
B88349
protein F15D4.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: B88349
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: B88349
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <STO>
A;Cross-references: UNIPROT:Q93511; UNIPARC:UPI0000179EBB; GB:chr_II; PIDN:CAB02486.1; PJ
C;Genetics:
A;Gene: F15D4.3
A;Map position: 2

Query Match 60.0%; Score 36; DB 2; Length 162;
Best Local Similarity 50.0%; Pred. No. 8.8;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLLQQVRVFV 12
|:|:|:|:|
Db 90 FMGVAQGLRYIF 101

RESULT 12
AH2090
two-component hybrid sensor and regulator alr2279 [imported] - Nostoc sp. (strain PCC 712
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2090
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2090
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1299 <KUR>
A;Cross-references: UNIPROT:Q8YUQ6; UNIPARC:UPI00000CE3A6; GB:BA000019; PIDN:BA873978.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr2279

Query Match 60.0%; Score 36; DB 2; Length 1299;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLLQQVRVFV 12
|:|:|:|:|
Db 71 FEGICQQTSLF 82

RESULT 13

G81105
Hypothetical protein NMB1242 [imported] - Neisseria meningitidis (strain MC58 serogroup C)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: G81105
R:Teketlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickel, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: G81105
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <TET>
A:Cross-references: UNIPROT:Q9JZ87; UNIPARC:UPI00000C4675; GB:AE002472; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1242
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1242

Query Match	58.3%	Score 35;	DB 2;	Length 102;	
Best Local Similarity	66.7%	Pred. No. 8.6;			
Matches	6;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

Qy 2 QGVVQQVRFV 10
Db 54 QAVMQQIRF 62

RESULT 14
S51728
NSP3 protein - human rotavirus (strain 69M)
N:Alternate names: NS34 protein
C:Species: human rotavirus
A:Variety: strain 69M
C>Date: 07-May-1995 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
R:Rao, C.D.; Das, M.; Rao, B.S.; Gowda, K.
submitted to the EMBL Data Library, September 1994
C:Accession: S51728
A:Reference number: S51709
A:Accession: S51728
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-313 <RAO>
A:Cross-references: UNIPROT:Q82051; UNIPARC:UPI00000F1AE7; EMBL:X81425; NID:g607091; PID
C:Superfamily: rotavirus nonstructural protein

Query Match	58.3%	Score 35;	DB 2;	Length 313;	
Best Local Similarity	41.7%	Pred. No. 29;			
Matches	5;	Conservative 5;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 1 FQGVVQQVRFV 12
Db 299 FKGLLQSQSNYYI 310

RESULT 15
S45068
53K glycoprotein - chicken
C:Species: Gallus gallus (chicken)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S45068
R:Boyd, K.L.; Kutchai, H.; Takeyasu, K.
submitted to the EMBL Data Library, January 1993
A:Description: Molecular cloning of the 53kDa glycoprotein from chicken sarcoplasmic reticulum
A:Reference number: S45068
A:Accession: S45068
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <BOY>
A:Cross-references: UNIPROT:Q90577; UNIPARC:UPI00000FE0AA; EMBL:Z21720; NID:g496324; PID

C:Superfamily: sarcoplasmic reticulum 53K glycoprotein
C:Keywords: glycoprotein

Query Match	58.3%	Score 35;	DB 2;	Length 471;	
Best Local Similarity	50.0%	Pred. No. 44;			
Matches	6;	Conservative 4;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 1 FQGVVQQVRFV 12
Db 52 FSGVQLRLKTY 63

Search completed: June 5, 2006, 22:45:00
Job time : 14.9655 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 108.931 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-20

Perfect score: 60

Sequence: 1 FQGVLOQVRFPV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: UniProt_sprot.*

2: UniProt_Trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	91.7	229	2	Q28194 BOVIN
2	55	91.7	496	2	Q7SY84_XENLA
3	55	91.7	1170	1	TSP1 BOVIN
4	55	91.7	1170	1	TSP1 HUMAN
5	55	91.7	1170	1	TSP1 MOUSE
6	55	91.7	1170	2	Q3TR40 MOUSE
7	55	91.7	1170	2	Q71SA3 RAT
8	55	91.7	1171	2	Q80YQ1 MOUSE
9	55	91.7	1171	2	Q8CGB2 MOUSE
10	55	91.7	1173	1	TSP1 XENLA
11	55	91.7	1225	2	Q59E59 HUMAN
12	50	83.3	1090	2	Q5SPG5 BRARE
13	50	83.3	1193	2	Q4S758 TETNG
14	49	81.7	249	2	Q5U903 PIG
15	49	81.7	1171	2	Q4RLR5 TETNG
16	45	75.0	1034	2	Q4RQ74 TETNG
17	41	68.3	367	2	Q6CKQ2 KLJLA
18	40	66.7	493	2	Q4CY08 TRYCR
19	40	66.7	1006	2	Q4DZ16 TRYCR
20	39	65.0	146	2	Q976Q1 SULFOB
21	39	65.0	214	2	Q82ZE3 ENTFA
22	39	65.0	313	2	Q4SHB8 TETNG
23	39	65.0	737	2	Q61R63 XENLA
24	39	65.0	780	2	Q5DJJ9 XENLA
25	39	65.0	1254	2	Q6CH81 YARLI
26	38	63.3	29	2	Q4XGT3 PLACH
27	38	63.3	212	1	HIS5_NEIMA
28	38	63.3	212	1	Q5FA21 NEIG1
29	38	63.3	327	2	Q5L310 GEOKA
30	38	63.3	469	2	Q3TZA3 CAREL
31	38	63.3	486	2	Q3JDG4_NITOC

32	38	63.3	494	2	Q5AVQ8 EMENI	Q5AVQ8 aspergillus
33	38	63.3	1168	2	Q5VH52_CIOIN	Q5VH52 ciona intes
34	37	61.7	210	2	Q48H98_PSE14	Q48H98 pseudomonas
35	37	61.7	210	2	Q4ZRP4_PSEU2	Q4ZRP4 pseudomonas
36	37	61.7	210	2	Q87ZW0_PSESM	Q87ZW0 pseudomonas
37	37	61.7	226	2	Q22GJ9_CALSA	Q22GJ9 caldicellul
38	37	61.7	318	1	TRUB_ACIAD	Q6F715 acinetobact
39	37	61.7	341	2	Q3FI58_9BURK	Q3FI58 burkholderi
40	37	61.7	494	2	Q2PFI9_9LAMI	Q2PFI9 verbena x h
41	37	61.7	597	2	Q4WLS2_ASPFU	Q4WLS2 aspergillus
42	37	61.7	889	2	Q5SD43_DICDI	Q5SD43 dictyosteli
43	37	61.7	1172	1	TSP2_HUMAN	P35442 homo sapien
44	37	61.7	1172	1	TSP2_MOUSE	Q03350 mus musculu
45	37	61.7	1172	2	Q5RI52_HUMAN	Q5RI52 homo sapien

ALIGNMENTS

RESULT 1
ID Q28194 BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96331130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCP19>3.3.CO;2-0;
RA Lafauillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
J. Cell. Physiol. 167:164-172(1996).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON_TER 1
FT NON_TER 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E8B669C CRC64;

Query Match 91.7%; Score 55; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.016;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRFPV 12

Db 190 FQGVLOQVRFPV 201

RESULT 2

Q7SY84_XENLA
ID Q7SY84_XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.

DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX Klein S., Strausberg R.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC054970; AAH54970.1; -; mRNA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR001007; VWFC.
CC Pfam; PF00090; TSP_1; 2.
CC Pfam; PF00093; VWC_1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 2.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS50092; TSP1; 2.
CC PROSITE; PS01208; VWFC_1; UNKNOWN_1.
CC PROSITE; PS0184; VWFC_2; 1.
SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7BF51B CRC64;

Query Match 91.7%; Score 55; DB 2; Length 496;
Best Local Similarity 91.7%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FQGLQVRFVF 12
| | | | | | | | | | | |

Db 214 FQGLQVRFVF 225
RESULT 3
TSPL_BOVIN
ID TSPL_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniprotKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSPL;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
RC MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RX Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
RT odontoblasts and preodontoblasts."
RL Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3. May play a role in dentinogenesis
CC and/or maintenance of dentin and dental pulp.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- TISSUE SPECIFICITY: Odontoblasts.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB005287; BAA21115.1; -; mRNA.
CC EMBL; X87618; CAA60950.1; -; mRNA.
CC EMBL; X87619; CAA60951.1; -; mRNA.
CC PIR; S55501; S55501.
CC HSSP; P07996; ILSL.
CC SMR; Q28178; 549-1169.
CC GlycoSuiteDB; Q28178; -.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSP_C.
CC InterPro; IPR001007; VWFC.
CC Pfam; PF00008; EGF_1.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 13.

DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS01184; VWC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT MOTIF 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT CARBOHYD 1085 1085
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT DISULFID 519 531
FT DISULFID 551 562
FT DISULFID 556 572
FT DISULFID 575 586
FT DISULFID 592 608
FT DISULFID 599 617
FT DISULFID 620 644
FT DISULFID 650 663
FT DISULFID 657 676
FT DISULFID 678 689
FT DISULFID 705 713
FT DISULFID 718 738
FT DISULFID 754 774
FT DISULFID 777 797
FT DISULFID 813 833
FT DISULFID 836 856
FT DISULFID 874 894
FT DISULFID 910 930
FT DISULFID 946 1167
FT CONFLICT 805 805 S -> G (in Ref. 2).
SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.086; 1; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 1;
QY 1 FQGVLLQQRVRFV 12
DB 208 FQGVLLQNRVRFV 219
RESULT 4
TSP1 HUMAN
ID TSP1 HUMAN STANDARD; PRT; 1170 AA.
AC P07996; O15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DE 07-MAR-2006, entry version 78.
DE Thrombospondin-1 precursor.
OS Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologies with several different
RT proteins";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region.";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Franson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;

RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.,
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RL module."; J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RN THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RN of the anti-angiogenic type 1 domains of human thrombospondin-1."; J.
RN Biochemistry 41:14329-14339(2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RN Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RN hydrazide chemistry, and mass spectrometry."; J.
RN Proteome Res. 4:2070-2080(2005).
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04665; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAA61237.1; ALT SEQ; mRNA.
DR EMBL; J04835; AAA61178.1; -; Genomic_DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1Z44; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGF; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; Gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPRO13320; ConA_like_subgrp.
DR InterPro; IPRO06210; EGF.
DR InterPro; IPRO00742; EGF_3.
DR InterPro; IPRO01881; EGF_Ca_bd.
DR InterPro; IPRO06209; EGF_like.
DR InterPro; IPRO13032; EGF like reg.
DR InterPro; IPRO03129; Laminin_G_TSP_N.
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO08085; TSP1.
DR InterPro; IPRO03367; tsp_3.
DR InterPro; IPRO08859; TSP_C.
DR InterPro; IPRO01007; VWFC.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWFC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS050184; VWFC_2; 1.
DR 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO_0000035842.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...).
FT N-linked (GlcNAc...).
FT C-linked (Man).
FT /FTID=CAR_000205.
FT O-linked (Fuc...).
FT /FTID=CAR_000206.
FT C-linked (Man).
FT /FTID=CAR_000207.
FT C-linked (Man).
FT /FTID=CAR_000208.
FT O-linked (Fuc...).
FT /FTID=CAR_000209.
FT C-linked (Man).
FT /FTID=CAR_000210.
FT O-linked (Fuc...).
FT /FTID=CAR_000211.
FT N-linked (GlcNAc...).
FT N-linked (GlcNAc...).
FT Interchain (Probable).
FT Interchain (Probable).

Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.088;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVVF 12

Db 208 FQGVLOQVRVVF 219

```

RESULT 5
TSPI_MOUSE
ID TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 57.
DE Thrombospondin-1 precursor.
GN Name=Thb1; Synonyms=Tspl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seidlin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression
RT during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfi D., Devarajulu S., Franson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of
RT the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
RN [4]
RP PROTEIN SEQUENCE OF 19-37.
RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
RT "Expression and initial characterization of recombinant mouse
RT thrombospondin 1 and thrombospondin 3.";
RL FEBS Lett. 387:36-41(1996).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; M62470; AAA50611.1; -; Genomic DNA.
EMBL; M62450; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62451; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62452; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62453; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62454; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62455; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62456; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62457; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62458; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62459; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62460; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62461; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62462; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62463; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62464; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62465; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62466; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62467; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62468; AAA50611.1; JOINED; Genomic DNA.
EMBL; M62469; AAA50611.1; JOINED; Genomic DNA.
EMBL; M87276; AAA53063.1; -; mRNA.
EMBL; J05606; AAA40431.1; -; Genomic DNA.
EMBL; J05605; AAA40431.1; JOINED; Genomic DNA.
PIR; A40558; A40558.
HSP; P07996; ILSL.
SMR; F35441; 549-1169.
Ensembl; ENSMUSG00000040152; Mus musculus.
MGI; MGI:98737; Thb1.
GO; GO:0005615; C:extracellular space; IDA.
GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
InterPro; IPR013320; ConA_like_subgrp.
InterPro; IPR006210; EGF.
InterPro; IPR001881; EGF_Ca bd.
InterPro; IPR006209; EGF_like.
InterPro; IPR013032; EGF_like_reg.
InterPro; IPR003129; Laminin_G_TSP_N.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP 1.
InterPro; IPR003367; TSP 3.
InterPro; IPR008859; TSP_C.
InterPro; IPR001007; VWF_C.
Pfam; PF00008; EGF; 2.
Pfam; PF00090; TSP 1; 3.
Pfam; PF02412; TSP 3; 12.
Pfam; PF05735; TSP_C; 1.
Pfam; PF00093; VWC; 1.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00181; EGF; 3.
SMART; SM00209; TSP1; 3.
SMART; SM00210; TSPN; 1.
SMART; SM00214; VWC; 1.
PROSITE; PS00022; EGF_1; FALSE_NEG.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS00026; EGF_3; 2.
PROSITE; PS00092; TSP1; 3.
PROSITE; PS01208; VWF_C; 1.
PROSITE; PS0184; VWF_2; 1.
PROSITE; PS0184; VWF_2; 1.
Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
KW Glycoprotein; Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT Thrombospondin-1.
FT /FTId=PRO_0000035843.
FT TSP N-terminal.
FT WFCC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708

```

```

FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 270 Interchain (Probable).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 391 423 By similarity.
FT DISULFID 395 428 By similarity.
FT DISULFID 406 413 By similarity.
FT DISULFID 447 484 By similarity.
FT DISULFID 451 489 By similarity.
FT DISULFID 462 474 By similarity.
FT DISULFID 504 541 By similarity.
FT DISULFID 508 546 By similarity.
FT DISULFID 519 531 By similarity.
FT DISULFID 551 562 By similarity.
FT DISULFID 556 572 By similarity.
FT DISULFID 575 586 By similarity.
FT DISULFID 592 608 By similarity.
FT DISULFID 599 617 By similarity.
FT DISULFID 620 644 By similarity.
FT DISULFID 650 663 By similarity.
FT DISULFID 657 676 By similarity.
FT DISULFID 678 689 By similarity.
FT DISULFID 705 713 By similarity.
FT DISULFID 718 738 By similarity.
FT DISULFID 754 774 By similarity.
FT DISULFID 777 797 By similarity.
FT DISULFID 813 833 By similarity.
FT DISULFID 836 856 By similarity.
FT DISULFID 874 894 By similarity.
FT DISULFID 910 930 By similarity.
FT DISULFID 946 1167 By similarity.
FT CONFLICT 1025 1025 F -> L (in Ref. 2).
SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.098;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQVRVFV 12
Db 208 FQGVQLQVRVFV 219

RESULT 6
Q3TR40_MOUSE PRELIMINARY; PRT; 1170 AA.
AC Q3TR40;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 5.
DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DN cDNA:AS3005N06 product:thrombospondin 1, full insert sequence.
GN Name=Fhbl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
[1]
NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RL "High-efficiency full-length cDNA cloning.";
[2]
Methods Enzymol. 303:19-44 (1999).
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Hayashizaki Y.;
RL "Analysis of the mouse transcriptome based on functional annotation of

```

```

RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa T.,
RA Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.P., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RL "The transcriptional landscape of the mammalian genome.";
Science 309:1559-1563 (2005).
[3]
NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
(RIKEN Genome Network Core Team) and the PANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
[4]
NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schirnl L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reid J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Haehizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```


DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF_3.
DR SMART; SM00209; TSP1_3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1_3.
DR PROSITE; PS01208; VWF_C; 1; UNKNOWN_1.
DR PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred.No. 0.088;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLQVRFVF 12
| | | | | | | | | |
Db 208 FQGVLLQVRFVF 219

RESULT 8

ID Q80YQ1_MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RC NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted under the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; BC050917; AAH50917.1; -, mRNA.
DR HSHP; P07996; ILSL.
DR SMR; Q80YQ1; 835-1170.
DR Ensembl; ENSMUSG0000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000805; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF_3;
DR SMART; SM00209; TSP1_3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1_3.
DR PROSITE; PS01208; VWF_C; 1; UNKNOWN_1.
DR PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred.No. 0.088;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLQVRFVF 12
| | | | | | | | | |
Db 208 FQGVLLQVRFVF 219

RESULT 9

Q8CGB2_MOUSE
ID Q8CGB2_MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q8CGB2;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-length enriched library, clone:G930018021 product:thrombospondin 1, full insert sequence).
DE Name=Thbs1;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RP STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RN [1]
RC NUCLEOTIDE SEQUENCE.
RP STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor
 RC arose spontaneously;
 RC Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RX Carninci P., Hayashizaki Y.;
 RA "High-efficiency full-length cDNA cloning.";
 RT Methods Enzymol. 303:19-44(1999).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RC PubMed=16141072; DOI=10.1126/science.1112014;
 RX Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilminger L.G., Aidinis V., Allen J.E.,
 RA Ambesi-Impombato A., Awewiler K.W., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummelink L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
 RA Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanihi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RT Science 309:1559-1563(2005).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.

TISSUE=Mammary gland;
 RC PubMed=16141073; DOI=10.1126/science.1112009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 RG (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566(2005).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Oato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKensie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Iehli Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [8]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Carninci P., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RA "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RT Genome Res. 10:1617-1630(2000).
 RN [9]

```
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [10]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC042422; AAH42422.1; -; mRNA.
DR EMBL; AK145202; BAE26293.1; -; mRNA.
DR HSSP; P07996; 1LSL.

Query Match          91.7%; Score 55; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. No. 0.088;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLQQRVRFV 12
Db 208 FQGVLLQNRVRFV 219

RESULT 10
TSPI1 XENLA
ID -TSPI1 XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-JAN-2006, entry version 54.
DE Thrombospondin-1 precursor.
GN Name=thbs1; Synonyms=tepl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
[1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RT "Cloning, characterization and expression of thrombospondin-1 in
RT Xenopus laevis embryos.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -----
```

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
DR HSSP; P07996; 1LSL.
DR SMR; P35448; 552-1172.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; WFCC_1; 1.
DR PROSITE; PS01184; WFCC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1173 /FTid=PRO_0000035844.
FT DOMAIN 23 224 TSP N-terminal.
FT DOMAIN 319 376 WFCC.
FT DOMAIN 382 432 TSP type-1 1.
FT DOMAIN 438 493 TSP type-1 2.
FT DOMAIN 495 550 TSP type-1 3.
FT DOMAIN 550 590 EGF-like 1.
FT DOMAIN 591 648 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 649 693 EGF-like 3.
FT DOMAIN 726 761 TSP type-3 1.
FT DOMAIN 762 784 TSP type-3 2.
FT DOMAIN 821 843 TSP type-3 3.
FT DOMAIN 844 881 TSP type-3 4.
FT DOMAIN 882 917 TSP type-3 5.
FT DOMAIN 918 953 TSP type-3 6.
FT DOMAIN 954 1173 TSP type-3 7.
FT REGION 23 235 Heparin-binding (Potential).
FT MOTIF 929 931 Cell attachment site (Potential).
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 363 363 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 705 705 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 711 711 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1070 1070 N-linked (GlcNAc...) (Potential).
FT DISULFID 394 426 By similarity.
FT DISULFID 398 431 By similarity.
FT DISULFID 409 416 By similarity.
FT DISULFID 450 487 By similarity.
FT DISULFID 454 492 By similarity.
FT DISULFID 465 477 By similarity.
FT DISULFID 507 544 By similarity.
FT DISULFID 511 549 By similarity.
FT DISULFID 522 534 By similarity.
```

```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 620 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 91.7%; Score 55; DB 1; Length 1173;
Best Local Similarity 91.7%; Pred. No. 0.089;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLQQRVRFV 12
DB 211 FQGVLQNRVRFV 222

RESULT 11
Q59E99 HUMAN
ID Q59E99_HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Tissue-Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
FT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AB209912; BAD93149.1; -; mRNA.
CC SMR; Q59E99; 886-939, 889-1225.
CC Ensembl; ENSG00000137801; Homo sapiens.
CC DR GO; GO:0005576; C:extracellular region; IEA.
CC DR GO; GO:0005509; F:calcium ion binding; IEA.
CC DR GO; GO:0008201; F:heparin binding; IEA.
CC DR GO; GO:0005515; F:protein binding; IEA.
CC DR GO; GO:0005198; F:structural molecule activity; IEA.
CC DR GO; GO:0007155; F:cell adhesion; IEA.
CC DR InterPro; IPR013320; ConA like subgrp.
CC DR InterPro; IPR006210; EGF.
CC DR InterPro; IPR00742; EGF_3.
CC DR InterPro; IPR001881; EGF_Ca bd.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR013032; EGF like reg.
CC DR InterPro; IPR003129; Laminin_G_TSP_N.
CC DR InterPro; IPR000884; TSP1.
CC DR InterPro; IPR008085; TSP_1.
CC DR InterPro; IPR008859; TSP_C.
```

```
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PRO1705; TSPREPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9880B16E57157B12 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1225;
Best Local Similarity 91.7%; Pred. No. 0.093;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLQQRVRFV 12
DB 263 FQGVLQNRVRFV 274

RESULT 12
Q5SPG5 BRARE
ID Q5SPG5_BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC
CC FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (by similarity).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AL928866; CAI20599.1; -; Genomic DNA.
CC SMR; Q5SPG5; 751-804, 754-1089.
CC Ensembl; ENSDARG0000010785; Danio rerio.
CC DR GO; GO:0005576; C:extracellular region; IEA.
CC DR GO; GO:0005509; F:calcium ion binding; IEA.
CC DR GO; GO:0008201; F:heparin binding; IEA.
CC DR GO; GO:0005515; F:protein binding; IEA.
CC DR GO; GO:0005198; F:structural molecule activity; IEA.
CC DR GO; GO:0007155; F:cell adhesion; IEA.
CC DR InterPro; IPR013320; ConA like subgrp.
CC DR InterPro; IPR002048; EF hand_Ca bd.
CC DR InterPro; IPR006210; EGF.
CC DR InterPro; IPR000742; EGF_3.
CC DR InterPro; IPR001881; EGF_Ca bd.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR013032; EGF like reg.
CC DR InterPro; IPR003129; Laminin_G_TSP_N.
CC DR InterPro; IPR000884; TSP1.
CC DR InterPro; IPR008085; TSP_1.
```

```
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF_2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS00186; EGF_2; 1.
DR PROSITE; PS00186; EGF_2; 1.
DR PROSITE; PS00266; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON TER 1
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;

Query Match 83.3%; Score 50; DB 2; Length 1090;
Best Local Similarity 83.3%; Pred. No. 0.87;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVQLQVRVVF 12
Db 185 FMGVQLQVRVVF 196

RESULT 13
Q4S758_TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758;
ID Q4S758_TETNG PRELIMINARY; PRT; 1193 AA.
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Chromosome 14 SCAP14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00022976001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
[1]
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaut S., Jaffe D., Fisher S., Lutfalla G., Bessat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellan S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier J., Chappell C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
[2]
RN NUCLEOTIDE SEQUENCE.
RP Genoscope; Whitehead Institute Centre for Genome Research;
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
-----
```

```
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; CAAE01014723; CAG03524.1; -; Genomic_DNA.
DR SMR; Q4S758; 811-1148
DR GO; GO:000576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR013032; EGF_like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF_2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Cell adhesion.
FT NON TER 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 83.3%; Score 50; DB 2; Length 1193;
Best Local Similarity 83.3%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVQLQVRVVF 12
Db 187 FMGVQLQVRVVF 198

RESULT 14
Q5U903_PIG PRELIMINARY; PRT; 249 AA.
ID Q5U903_PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
NCBI_TaxID=9823;
[1]
RN NUCLEOTIDE SEQUENCE.
RA Zhang K., Mauc G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY773342; AAV38110.1; -; mRNA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
-----
```

```
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC_1; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS0184; VWC_2; 1.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 27560 MW; 465D664BE0329C32 CRC64;

Query Match 81.7%; Score 49; DB 2; Length 249;
Best Local Similarity 90.9%; Pred. No. 0.29; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVVLQVRVVF 12
DB 1 QGVVLQVRVVF 11

RESULT 15
Q4RLR5_TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, entry version 8.
DE Chromosome 10 SCAP15019, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00032374001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dessat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn W., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; CAAB01015019; CAG10667.1; -; Genomic_DNA.
DR SMR; Q4RLR5; 834-887, 837-1171.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
```

```
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWC_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS0184; VWC_2; 1.
KW Cell adhesion.
FT NON_TER 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 81.7%; Score 49; DB 2; Length 1171;
Best Local Similarity 83.3%; Pred. No. 1.5; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FGVLQVRVVF 12
DB 207 FGVLQVRVVF 218

Search completed: June 5, 2006, 22:42:40
Job time : 109.931 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-20

Perfect score: 60

Sequence: 1 FQGVLLQVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /EMC_Celerra_SID33/ptodata/2/iaa/5 COMB.pbp:*
2: /EMC_Celerra_SID33/ptodata/2/iaa/6 COMB.pbp:*
3: /EMC_Celerra_SID33/ptodata/2/iaa/7 COMB.pbp:*
4: /EMC_Celerra_SID33/ptodata/2/iaa/H COMB.pbp:*
5: /EMC_Celerra_SID33/ptodata/2/iaa/PCTUS COMB.pbp:*
6: /EMC_Celerra_SID33/ptodata/2/iaa/RE COMB.pbp:*
7: /EMC_Celerra_SID33/ptodata/2/iaa/backfiles1.pbp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	91.7	825	2	US-09-949-002-482
2	55	91.7	831	2	US-09-939-853A-97
3	55	91.7	831	2	US-09-939-853A-98
4	55	91.7	1170	1	US-08-313-288B-20
5	55	91.7	1170	2	US-09-657-472-2
6	55	91.7	1170	2	US-09-949-002-350
7	39	65.0	169	2	US-09-134-000C-3511
8	37	61.7	1045	2	US-09-949-016-1112
9	37	61.7	1172	1	US-08-313-288B-19
10	37	61.7	1172	2	US-09-949-016-6333
11	36	60.0	175	2	US-09-252-991A-18826
12	36	60.0	440	1	US-08-307-499-15
13	36	60.0	440	2	US-09-299-268-15
14	36	60.0	516	2	US-09-291-170A-1
15	36	60.0	516	2	US-09-724-884-1
16	36	60.0	517	2	US-09-724-592-1
17	36	60.0	517	2	US-09-673-222-1
18	35	58.3	1142	2	US-09-106-075A-89
19	35	58.3	1333	2	US-09-347-878-20
20	35	58.3	1360	2	US-09-949-016-7739
21	34	56.7	539	2	US-09-800-170-16
22	34	56.7	710	2	US-09-489-039A-14121
23	34	56.7	944	2	US-09-134-000C-5578
24	34	56.7	1019	1	US-08-271-364A-7
25	34	56.7	1019	1	US-08-222-715B-26
26	34	56.7	1300	2	US-09-698-341-3

```

27 33 55.0 92 2 US-09-902-540-13721 Sequence 13721, A
28 33 55.0 158 2 US-09-248-796A-17510 Sequence 17510, A
29 33 55.0 269 1 US-09-006-535-6 Sequence 6, Appli
30 33 55.0 287 2 US-09-258-634-5 Sequence 5, Appli
31 33 55.0 296 1 US-09-006-535-5 Sequence 5, Appli
32 33 55.0 296 2 US-09-949-016-6605 Sequence 6605, Ap
33 33 55.0 336 2 US-09-949-016-7481 Sequence 7481, Ap
34 33 55.0 344 2 US-09-328-352-6058 Sequence 6058, Ap
35 33 55.0 349 2 US-09-949-016-9668 Sequence 9668, Ap
36 33 55.0 350 2 US-09-902-540-11953 Sequence 11953, A
37 33 55.0 396 2 US-09-252-991A-28379 Sequence 28379, A
38 33 55.0 432 2 US-09-489-039A-7909 Sequence 7909, A
39 33 55.0 453 2 US-10-094-749-1719 Sequence 1719, Ap
40 33 55.0 510 2 US-09-252-991A-25258 Sequence 25258, A
41 33 55.0 582 2 US-09-270-767-45532 Sequence 45532, A
42 33 55.0 715 1 US-08-484-993B-10 Sequence 10, Appl
43 33 55.0 715 1 US-08-484-158B-10 Sequence 10, Appl
44 33 55.0 715 1 US-08-484-596A-10 Sequence 10, Appl
45 33 55.0 715 1 US-08-480-150A-10 Sequence 10, Appl

```

ALIGNMENTS

RESULT 1

```

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CLO00790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-482

```

```

Query Match 91.7%; Score 55; DB 2; Length 825;
Best Local Similarity 91.7%; Pred. No. 0.014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 FQGVLLQVRFVF 12
    |||||
Db 308 FQGVLLQVRFVF 319

```

RESULT 2

```

US-09-939-853A-97
; Sequence 87, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20

```

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.7%; Score 55; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQQVRVF 12
Db      208 FQGVQNVRFV 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232a1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          91.7%; Score 55; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQQVRVF 12
Db      208 FQGVQNVRFV 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and AviHu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQQVRVF 12
Db      208 FQGVQNVRFV 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQQVRVF 12
Db      208 FQGVQNVRFV 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-350

Query Match 91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLQVRVF 12
Db 208 FQGVLLQVRVF 219

RESULT 7
US-09-134-000C-3511
; Sequence 3511, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3511
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (29)-(70)
; OTHER INFORMATION: Amino acids 29, 30, 59 & 70 are Xaa wherein Xaa = any amino acid.
US-09-134-000C-3511

Query Match 65.0%; Score 39; DB 2; Length 169;
Best Local Similarity 77.8%; Pred. No. 4.2;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLLQVR 9
Db 76 FQGVLLQVR 84

RESULT 8
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11112

Query Match 61.7%; Score 37; DB 2; Length 1045;
Best Local Similarity 58.3%; Pred. No. 79;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLLQVRVF 12
Db 281 FQGVLLQVRVF 292

RESULT 9
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match 61.7%; Score 37; DB 1; Length 1172;
Best Local Similarity 58.3%; Pred. No. 89;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLLQVRVF 12
Db 202 FQGVLLQVRVF 213

RESULT 10

```
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
;
US-09-949-016-6333
;
Query Match 61.7%; Score 37; DB 2; Length 1172;
Best Local Similarity 58.3%; Pred. No. 89;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVQQVRFVF 12
Db 202 FRGLLNVLVF 213

RESULT 11
US-09-252-991A-18826
; Sequence 18826, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18826
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-18826
;
Query Match 60.0%; Score 36; DB 2; Length 175;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQQVRFVF 12
Db 41 QGVLDVQVFLF 51

RESULT 12
US-08-307-499-15
; Sequence 15, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;
US-09-299-268-15
; Sequence 15, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;
US-08-307-499-15
;
Query Match 60.0%; Score 36; DB 1; Length 440;
Best Local Similarity 41.7%; Pred. No. 48;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVQQVRFVF 12
Db 272 FKSIVQVVFIF 283

RESULT 13
US-09-299-268-15
; Sequence 15, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;
US-08-307-499-15
;
Query Match 60.0%; Score 36; DB 1; Length 440;
Best Local Similarity 41.7%; Pred. No. 48;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVQQVRFVF 12
Db 272 FKSIVQVVFIF 283

TITLE OF INVENTION: Live Vaccine Vector
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,499
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-307-499-15
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-268-15

Query Match 60.0%; Score 36; DB 2; Length 440;
Best Local Similarity 41.7%; Pred. No. 48;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRVF 12
Db 272 FKSIVQYVKFIF 283

RESULT 14
US-09-291-170A-1
Sequence 1, Application US/09291170A
Patent No. 6410697
GENERAL INFORMATION:
APPLICANT: Vale, Ronald D.
APPLICANT: Hartman, James J.
TITLE OF INVENTION: Assays for the Detection of Microtubule
TITLE OF INVENTION: Depolymerization Inhibitors
FILE REFERENCE: 18557B-000510US
CURRENT APPLICATION NUMBER: US/09/291,170A
CURRENT FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 60/081,734
PRIOR FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 516
TYPE: PRT
ORGANISM: Strongylocentrotus purpuratus
FEATURE:
OTHER INFORMATION: katanin p60 subunit
US-09-291-170A-1

Query Match 60.0%; Score 36; DB 2; Length 516;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLOQVR 9
Db 29 YQGVLOQIQ 37
RESULT 15
US-09-724-884-1
Sequence 1, Application US/09724884
Patent No. 6429304
GENERAL INFORMATION:
APPLICANT: Vale, Ronald D.
APPLICANT: Hartman, James J.
TITLE OF INVENTION: Assays for the Detection of Microtubule
TITLE OF INVENTION: Depolymerization Inhibitors
FILE REFERENCE: 18557B-000510US
CURRENT APPLICATION NUMBER: US/09/724,884
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 09/291,170
PRIOR FILING DATE: 1999-04-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 516
TYPE: PRT
ORGANISM: Strongylocentrotus purpuratus
FEATURE:
OTHER INFORMATION: katanin p60 subunit
US-09-724-884-1

Query Match 60.0%; Score 36; DB 2; Length 516;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLOQVR 9
Db 29 YQGVLOQIQ 37

Search completed: June 5, 2006, 22:48:54
Job time : 24.8966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-20

Perfect score: 60

Sequence: 1 FQGVQQVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	55	91.7	12	4	US-10-474-213-28		Sequence 28, Appl
2	55	91.7	240	4	US-10-419-462-40		Sequence 40, Appl
3	55	91.7	240	5	US-10-782-968-40		Sequence 40, Appl
4	55	91.7	432	5	US-10-741-600-1020		Sequence 1020, Ap
5	55	91.7	432	5	US-10-741-600-1022		Sequence 1022, Ap
6	55	91.7	459	6	US-11-043-806-462		Sequence 462, App
7	55	91.7	466	3	US-09-925-301-1047		Sequence 1047, Ap
8	55	91.7	555	6	US-11-043-806-454		Sequence 454, App
9	55	91.7	578	6	US-11-043-806-456		Sequence 456, App
10	55	91.7	685	6	US-11-043-806-452		Sequence 452, App
11	55	91.7	804	6	US-11-043-806-453		Sequence 453, App
12	55	91.7	828	6	US-11-043-806-455		Sequence 455, App
13	55	91.7	831	3	US-09-939-853A-97		Sequence 97, Appl
14	55	91.7	831	3	US-09-939-853A-98		Sequence 98, Appl
15	55	91.7	855	6	US-11-043-806-461		Sequence 461, App
16	55	91.7	1000	6	US-11-043-806-457		Sequence 457, App
17	55	91.7	1105	6	US-11-043-806-458		Sequence 458, App
18	55	91.7	1150	4	US-10-296-733-1		Sequence 1, Appli
19	55	91.7	1152	3	US-09-919-603-1		Sequence 1, Appli
20	55	91.7	1169	5	US-10-317-821B-7		Sequence 7, Appli
21	55	91.7	1170	4	US-10-020-141-12		Sequence 12, Appl
22	55	91.7	1170	4	US-10-017-721-2		Sequence 2, Appli
23	55	91.7	1170	4	US-10-021-660-114		Sequence 114, App
24	55	91.7	1170	4	US-10-008-093-2		Sequence 2, Appli
25	55	91.7	1170	4	US-10-295-027-1170		Sequence 1170, Ap
26	55	91.7	1170	4	US-10-211-462-38		Sequence 38, Appl
27	55	91.7	1170	4	US-10-231-956A-482		Sequence 482, App

28	55	91.7	1170	4	US-10-419-462-38	Sequence 38, Appl
29	55	91.7	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	55	91.7	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	55	91.7	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	55	91.7	1170	5	US-10-782-968-38	Sequence 38, Appl
33	55	91.7	1170	5	US-10-849-989-44	Sequence 44, Appl
34	55	91.7	1170	5	US-10-631-467-548	Sequence 548, App
35	55	91.7	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	55	91.7	1170	5	US-10-831-997-2	Sequence 2, Appli
37	55	91.7	1170	5	US-10-995-561-594	Sequence 594, App
38	55	91.7	1170	5	US-10-995-561-595	Sequence 595, App
39	55	91.7	1170	5	US-10-995-561-596	Sequence 596, App
40	55	91.7	1170	6	US-11-037-713-51	Sequence 51, Appl
41	55	91.7	1170	6	US-11-046-644-28	Sequence 28, Appl
42	55	91.7	1170	6	US-11-046-456-28	Sequence 28, Appl
43	39	65.0	15	4	US-10-285-394-153	Sequence 153, App
44	38	63.3	226	5	US-10-467-657-2428	Sequence 2428, Ap
45	37	61.7	16	3	US-09-822-682-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Krutzsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/105335
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQQVRFVF 12
DB 1 FQGVQQVRFVF 12

RESULT 2
US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J
; APPLICANT: Williams, Kevin J
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40

```
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match          91.7%; Score 55; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLQQRVRFVF 12
Db 190 FQGVLQNRVRFVF 201

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match          91.7%; Score 55; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLQQRVRFVF 12
Db 190 FQGVLQNRVRFVF 201

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match          91.7%; Score 55; DB 5; Length 432;
```

```
Best Local Similarity 91.7%; Pred. No. 0.095;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLQQRVRFVF 12
Db 208 FQGVLQNRVRFVF 219

RESULT 5
US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match          91.7%; Score 55; DB 5; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.095;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLQQRVRFVF 12
Db 208 FQGVLQNRVRFVF 219

RESULT 6
US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match          91.7%; Score 55; DB 6; Length 459;
Best Local Similarity 91.7%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLQQRVRFVF 12
Db 208 FQGVLQNRVRFVF 219

RESULT 7
US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```


; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 91.7%; Score 55; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRFFV 12
| | | | | | | | | |
Db 261 FQGVLOQVRFFV 272

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 91.7%; Score 55; DB 6; Length 555;
Best Local Similarity 91.7%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRFFV 12
| | | | | | | | | |
Db 208 FQGVLOQVRFFV 219

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 91.7%; Score 55; DB 6; Length 578;
Best Local Similarity 91.7%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRFFV 12
| | | | | | | | | |
Db 208 FQGVLOQVRFFV 219

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 91.7%; Score 55; DB 6; Length 685;
Best Local Similarity 91.7%; Pred. No. 0.15;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRFFV 12
| | | | | | | | | |
Db 208 FQGVLOQVRFFV 219

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 91.7%; Score 55; DB 6; Length 804;
Best Local Similarity 91.7%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLOQVRFFV 12
| | | | | | | | | |
Db 208 FQGVLOQVRFFV 219

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455
Query Match 91.7%; Score 55; DB 6; Length 828;
Best Local Similarity 91.7%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FQGVLOQVRVVF 12
Db 208 FQGVLOQVRVVF 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match 91.7%; Score 55; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FQGVLOQVRVVF 12
Db 208 FQGVLOQVRVVF 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus

US-09-939-853A-98

Query Match 91.7%; Score 55; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FQGVLOQVRVVF 12
Db 208 FQGVLOQVRVVF 219

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847,1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match 91.7%; Score 55; DB 6; Length 855;
Best Local Similarity 91.7%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FQGVLOQVRVVF 12
Db 208 FQGVLOQVRVVF 219

Search completed: June 6, 2006, 00:00:10
Job time : 78.6207 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-20

Perfect score: 60

Sequence: 1 FQGVLLQVRFVF 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	58.3	1333	6	US-10-511-937-2992
2	34	56.7	1339	7	US-11-293-697-3323
3	34	56.7	490	7	US-11-293-697-4335
4	33	55.0	164	6	US-10-953-349-1159
5	33	55.0	249	6	US-10-953-349-1158
6	33	55.0	250	6	US-10-953-349-1157
7	33	55.0	287	6	US-10-953-349-1067
8	33	55.0	315	6	US-10-953-349-9204
9	33	55.0	340	6	US-10-953-349-1066
10	33	55.0	400	6	US-10-953-349-9203
11	33	55.0	401	6	US-10-953-349-9202
12	33	55.0	609	7	US-11-293-697-3371
13	32	53.3	763	7	US-11-293-697-3377
14	32	53.3	1504	6	US-10-505-928-662
15	31	51.7	290	6	US-10-953-349-5462
16	31	51.7	298	6	US-10-953-349-9347
17	31	51.7	329	6	US-10-953-349-5461
18	31	51.7	341	7	US-11-317-983-8
19	31	51.7	503	7	US-11-317-983-9
20	31	51.7	806	7	US-11-293-697-4195
21	30	50.0	186	6	US-10-953-349-27518
22	30	50.0	188	6	US-10-953-349-27517
23	30	50.0	213	6	US-10-953-349-27516
24	30	50.0	318	6	US-10-953-349-987
25	30	50.0	361	6	US-10-953-349-986

26	30	50.0	386	6	US-10-953-349-985	Sequence 985, App
27	30	50.0	525	6	US-10-953-349-31207	Sequence 31207, A
28	30	50.0	560	6	US-10-953-349-31206	Sequence 31206, A
29	30	50.0	820	6	US-10-982-908-26	Sequence 26, Appl
30	29	48.3	90	6	US-10-953-349-24350	Sequence 24350, A
31	29	48.3	332	6	US-10-953-349-3565	Sequence 3565, A
32	29	48.3	340	7	US-11-293-697-3628	Sequence 3628, Ap
33	29	48.3	369	6	US-10-953-349-9207	Sequence 9207, Ap
34	29	48.3	391	6	US-10-953-349-21132	Sequence 21132, A
35	29	48.3	403	6	US-10-953-349-21131	Sequence 21131, A
36	29	48.3	412	6	US-10-953-349-9206	Sequence 9206, Ap
37	29	48.3	428	6	US-10-953-349-21130	Sequence 21130, A
38	29	48.3	438	6	US-10-953-349-28462	Sequence 28462, A
39	29	48.3	442	6	US-10-953-349-28461	Sequence 28461, A
40	29	48.3	446	6	US-10-713-648A-57	Sequence 57, Appl
41	29	48.3	453	6	US-10-953-349-28460	Sequence 28460, A
42	29	48.3	455	6	US-10-953-349-35597	Sequence 35597, A
43	29	48.3	474	6	US-10-953-349-13546	Sequence 13546, A
44	29	48.3	482	6	US-10-953-349-13545	Sequence 13545, A
45	29	48.3	494	6	US-10-953-349-35596	Sequence 35596, A

ALIGNMENTS

RESULT 1
US-10-511-937-2992
; Sequence 2992, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2992
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2992

Query Match 58.3%; Score 35; DB 6; Length 1333;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLLQVRF 10
|:|:|:|:|:
Db 327 FRGVLEQLRW 336

RESULT 2
US-11-293-697-3323
; Sequence 3323, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA

; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3323
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3323

Query Match 56.7%; Score 34; DB 7; Length 139;
Best Local Similarity 66.7%; Pred. No. 4.6;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQQRVF 11
| : ||| :
Db 89 GSMQQRVF 97

RESULT 3

US-11-293-697-4335
; Sequence 4335, Application US/11/293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4335
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4335

Query Match 56.7%; Score 34; DB 7; Length 490;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQQVR 9
| : ||| : ||| :
Db 29 YQGVQQIQ 37

RESULT 4

US-10-953-349-1159
; Sequence 1159, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1159
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1159

Query Match 55.0%; Score 33; DB 6; Length 164;
Best Local Similarity 50.0%; Pred. No. 8.7;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVQQVRVF 12
| : | : ||| :
Db 82 FEAVVDRVRLVF 93

RESULT 5

US-10-953-349-1158
; Sequence 1158, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1158
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1158

Query Match 55.0%; Score 33; DB 6; Length 249;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVQQVRVF 12
| : | : ||| :
Db 167 FEAVVDRVRLVF 178

RESULT 6

US-10-953-349-1157
; Sequence 1157, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1157
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1157

Query Match 55.0%; Score 33; DB 6; Length 250;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVQQVRVF 12
| : | : ||| :
Db 168 FEAVVDRVRLVF 179

RESULT 7

US-10-953-349-1067
; Sequence 1067, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1067
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1067

Query Match 55.0%; Score 33; DB 6; Length 287;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLQQVRFFV 11
Db 11 YQSSLSQIRFV 21

RESULT 8

US-10-953-349-9204
; Sequence 9204, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9204
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9204

Query Match 55.0%; Score 33; DB 6; Length 315;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLQQVRFFV 12
Db 82 FEAVVDRVRLVF 93

RESULT 9

US-10-953-349-1066
; Sequence 1066, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1066
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1066

Query Match 55.0%; Score 33; DB 6; Length 340;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLQQVRFFV 11
Db 64 YQSSLSQIRFV 74

RESULT 10

US-10-953-349-9203
; Sequence 9203, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9203
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9203

Query Match 55.0%; Score 33; DB 6; Length 400;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLQQVRFFV 12
Db 167 FEAVVDRVRLVF 178

RESULT 11

US-10-953-349-9202
; Sequence 9202, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9202
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9202

Query Match 55.0%; Score 33; DB 6; Length 401;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLQQVRFFV 12
Db 168 FEAVVDRVRLVF 179

RESULT 12

US-11-293-697-3971
; Sequence 3971, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3971
; LENGTH: 609
; TYPE: PRT

```
; ORGANISM: Homo sapiens
US-11-293-697-3971

Query Match      55.0%; Score 33; DB 7; Length 609;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy      1 FQGVLOQVRVVF 12
       |:|:|:|:|
Db      6 FQGLITQARLPF 17

RESULT 13
US-11-293-697-3277
; Sequence 3277, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3277
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3277

Query Match      53.3%; Score 32; DB 7; Length 763;
Best Local Similarity 75.0%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QGVLOQVR 9
       |:|:|:|
Db      389 EGLLOQVR 396

RESULT 14
US-10-505-928-662
; Sequence 662, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 662
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-662

Query Match      53.3%; Score 32; DB 6; Length 1504;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 QGVLOQVRVVF 12
       |:|:|:|
Db      751 QHILSLRPFV 761

RESULT 15
US-10-953-349-5462
; Sequence 5462, Application US/10953349
```

```
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5462
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5462

Query Match      51.7%; Score 31; DB 6; Length 290;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      3 GVLQQVRVVF 12
       |:|:|:|
Db      142 GLLEEVREVF 151

Search completed: June 6, 2006, 00:12:56
Job time : 4.82414 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-21

Perfect score: 59

Sequence: 1 FQGVLSVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq 8:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*
- 9: Geneseqp2005s:*
- 10: Geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	59	100.0	12	4	AAB35376	Aab35376 Alpha3bet
2	56	94.9	12	4	AAB35352	Aab35352 Alpha3bet
3	56	94.9	12	4	AAB35361	Aab35361 Alpha3bet
4	56	94.9	12	4	AAB35378	Aab35378 Alpha3bet
5	56	94.9	12	6	ABG72834	Abg72834 Thrombos
6	56	94.9	240	8	ADL70641	Adl70641 Human thr
7	56	94.9	432	8	ADQ39359	Adq39359 Human myo
8	56	94.9	432	8	ADQ39357	Adq39357 Human myo
9	56	94.9	459	4	AAU02916	Aau02916 Angiotens
10	56	94.9	466	3	AB43602	Ab43602 Human can
11	56	94.9	546	4	AAU02915	Aau02915 Angiotens
12	56	94.9	548	7	ADN02474	Adn02474 TSF polyp
13	56	94.9	555	4	AAU02914	Aau02914 Angiotens
14	56	94.9	731	3	AAU02913	Aau02913 Angiotens
15	56	94.9	1152	3	AAB00042	Aab00042 Human thr
16	56	94.9	1152	5	AAU74771	Aau74771 Human thr
17	56	94.9	1152	5	ABB82285	Abb82285 Human thr
18	56	94.9	1170	4	AB74450	Ab74450 Human var
19	56	94.9	1170	4	AB90800	Ab90800 Human she
20	56	94.9	1170	5	AAE25030	Aae25030 Human thr
21	56	94.9	1170	5	AAU75315	Aau75315 Human thr
22	56	94.9	1170	6	ABP96780	Abp96780 Human COP
23	56	94.9	1170	6	ABU03474	Abu03474 Angiogene

24	56	94.9	1170	6	ABG74673	Abg74673 Human THB
25	56	94.9	1170	6	AAE36228	Aae36228 Human THB
26	56	94.9	1170	7	ABR62059	AbR62059 Human thr
27	56	94.9	1170	7	ADN39852	Adn39852 Cancer/an
28	56	94.9	1170	8	ADJ76124	Adj76124 Marker ge
29	56	94.9	1170	8	ADJ75296	Adj75296 Marker ge
30	56	94.9	1170	8	ADL70639	Adl70639 Human thr
31	56	94.9	1170	8	ADL35874	Adl35874 Human thr
32	56	94.9	1170	8	ADQ26070	Adq26070 Thrombos
33	56	94.9	1170	8	ADP54179	Adp54179 Human PRO
34	56	94.9	1170	8	ADQ39358	Adq39358 Human myo
35	56	94.9	1170	8	ADQ39356	Adq39356 Human myo
36	56	94.9	1170	8	ADQ39355	Adq39355 Human myo
37	56	94.9	1170	9	AD221688	Ad221688 Thrombos
38	56	94.9	1170	9	ABE87781	Aeb87781 Human thr
39	56	94.9	1170	9	ABE46751	Aeb46751 Human thr
40	55	93.2	12	4	AAB35379	Aab35379 Alpha3bet
41	55	93.2	12	4	AAB35380	Aab35380 Alpha3bet
42	53	89.8	12	4	AAB35373	Aab35373 Alpha3bet
43	53	89.8	12	4	AAB35381	Aab35381 Alpha3bet
44	52	88.1	12	4	AAB35364	Aab35364 Alpha3bet
45	52	88.1	12	4	AAB35374	Aab35374 Alpha3bet

ALIGNMENTS

RESULT 1
AAB35376
ID AAB35376 standard; peptide; 12 AA.
AC AAB35376;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3bet integrin binding peptide #41.
XX
KW Alpha3bet integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Roberts DD, Krutzsch HC;
XX WPI; 2001-182656/18.
XX
XX New peptides that bind to or are recognized by alpha3-beta1 integrins,
XX useful for inhibiting cell adhesion to extracellular matrix, cell
XX motility and proliferation and for treating rheumatoid arthritis and
XX cancer.
XX Claim 4; Page 34; 84pp; English.
XX
XX The present invention provides a number of peptides which bind to
XX alpha3beta1 integrins. They are useful in the modulation of cell adhesion
XX and motility, and in the treatment of cancer, diabetic retinopathy,
XX rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
XX of the invention
XX Sequence 12 AA;

```
Query Match      100.0%; Score 59; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
Db 1 FQGVLSQVRVVF 12

RESULT 2
AAB35352
ID AAB35352 standard; peptide; 12 AA.
XX
AC AAB35352;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #17.
XX
KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Example 2; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;

Query Match      94.9%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00044;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
Db 1 FQGVLSQVRVVF 12

RESULT 3
AAB35361
ID AAB35361 standard; peptide; 12 AA.
XX
AC AAB35361;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #26.
XX
```

```
XX Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Example 2; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;

Query Match      94.9%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00044;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
Db 1 FQGVLSQVRVVF 12

RESULT 4
AAB35378
ID AAB35378 standard; peptide; 12 AA.
XX
AC AAB35378;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #43.
XX
KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
```


XX WPI; 2001-182656/18.
XX New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX Example 2; Page 34; 84pp; English.
PS
XX The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;
Query Match 94.9%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00044;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 FQGVLSQVRVFV 12
Db 1 FQGVLSQVRVFV 12
RESULT 5
ID ABG72834 standard; peptide; 12 AA.
XX
AC ABG72834;
XX
DT 24-FEB-2003 (first entry)
XX
DE Thrombospondin-1 sequence containing synthetic peptide.
XX
KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
OS Synthetic.
XX
XX WO200281630-A2.
XX
XX 17-OCT-2002.
XX
XX 03-APR-2002; 2002WO-US010535.
XX
XX 06-APR-2001; 2001US-0281994P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Roberts DD, Krutzsch HC;
XX
XX WPI; 2003-103329/09.
XX
XX A new diagnosis for cancer other than prostate cancer in a mammal useful
PT to detect cancer including lung cancer, particularly small cell lung
PT cancer and melanoma comprises detecting semenogelin in a sample.
XX
XX Example 1; Page 14; 32pp; English.
XX
XX The invention relates to diagnosing cancer other than prostate cancer in
CC a male mammal, comprising assaying a test sample for increased level of
CC semenogelin, or cancer in a female by assaying for the presence of
CC semenogelin. Administering a semenogelin protein or polypeptide fragment
CC or a semenogelin-specific antibody or active fragment, or a recombinant
CC vector expressing the protein or antibody, is useful for inducing an
CC immune response to a cancer in a mammal, where the cancer is not prostate
CC cancer and semenogelin is a marker. The invention is used to diagnose
CC cancer, particularly of epithelial origin such as lung cancer, papillary

CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
CC (SCLC), or a melanoma. The present sequence represents the amino acid
CC sequence of the thrombospondin-1 sequence containing synthetic peptide
CC which binds to alpha-3-beta-1 integrin
XX
SQ Sequence 12 AA;
Query Match 94.9%; Score 56; DB 6; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00044;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 FQGVLSQVRVFV 12
Db 1 FQGVLSQVRVFV 12
RESULT 6
ID ADL70641 standard; protein; 240 AA.
XX
AC ADL70641;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human thrombospondin-1 N-terminal domain.
XX
KW Human; thrombospondin-1; epitope; cancer; diagnosis.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 23..32
FT /note= "Heparin binding region"
FT Region 77..82
FT /note= "Heparin binding region"
FT Region 151..164
FT /note= "Fibrinogen binding region"
XX
XX WO2004018995-A2.
XX
XX 04-MAR-2004.
XX
XX 20-AUG-2003; 2003WO-US026023.
XX
XX 23-AUG-2002; 2002US-0405494P.
XX
XX 21-APR-2003; 2003US-00419462.
XX
XX (WILL/) WILLIAMS K J.
XX
XX Williams KJ;
XX
XX WPI; 2004-226901/21.
XX
XX New purified thrombospondin fragment extracted from a body fluid, useful
PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
PT or leukemia or as calibrators, indicators, immunogens and analytes.
XX
XX Disclosure; SEQ ID NO 40; 76pp; English.
XX
XX The present sequence is that of the N-terminal domain of human
CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
CC clinical assays for cancer and for generation of antibodies and other
CC binding agents. A method that distinguishes TSP from a TSP fragment or
CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
CC or portion as a target for a binding molecule, e.g. an antibody, to
CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
CC epitope present in TSP but not in the fragment or portion to obtain a
CC quantitation of TSP only; and (3) using the difference between (1) and
CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
CC the TSP fragment or portion is performed in order to detect the presence,
CC or monitor the course, of a disease or condition selected from cancer,

renal failure, renal disease, atopic dermatitis, vasculitis, acute
vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
thrombocytopenic purpura, purpura, an inflammatory condition, a condition
associated with clotting, a condition associated with platelet
activation, a condition associated with intravascular platelet
activation, a condition associated with consumption of platelets, heparin
-induced thrombocytopenia, disseminated intravascular coagulation,
intravascular coagulation, extravascular coagulation, a condition
associated with endothelial activation, a condition associated with
production and/or release of thrombospondin and/or a thrombospondin
fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
reaction, an aspartame reaction, atopic dermatitis, eczema,
hypersensitivity, scleroderma, conditions associated with plugging of
vessels, a condition associated with a cryofibrinogen, a condition
associated with a cryoglobulin, and a condition associated with an anti-
cardiolipin antibody. The cancer is selected from adenoma,
adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
a cancer with vascular invasion, internal cancer, skin cancer, cancer of
the respiratory system, circulatory system, musculoskeletal system,
muscle, bone, a joint, tendon or ligament, digestive system, liver or
biliary system, pancreas, head, neck, endocrine system, reproductive
system (male or female), genitourinary system, kidney, urinary tract,
sensory system, nervous system, lymphoid organ, blood, a gland, mammary
gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
well-differentiated cancer or a moderately differentiated cancer.

SQ Sequence 240 AA;

Query Match 94.9%; Score 56; DB 8; Length 240;

Best Local Similarity 91.7%; Pred. No. 0.012;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12

| | | | | : | | | | |

Db 190 FQGVLSVRVFV 201

RESULT 7

ADQ39359

ID ADQ39359 standard; protein; 432 AA.

XX ADQ39359;

XX ADQ39359;

XX ADQ39359;

DT 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

DE Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

OS Homo sapiens.

XX Homo sapiens.

XX WO2004058052-A2.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

DR WPI; 2004-533949/51.

XX N-PSDB; ADQ38531.

XX Identifying an individual who has an altered risk for developing

PT myocardial infarction by detecting a single nucleotide polymorphism in

PT the individual's nucleic acids.

XX Claim 10; SEQ ID NO 1022; 145pp; English.

XX The invention relates to a novel method for identifying an individual who

CC has an altered risk for developing myocardial infarction. The method

CC comprises detecting a single nucleotide polymorphism (SNP) in any one of

CC the nucleotide sequences given in the specification in the individual's

CC nucleic acids, where the presence of the SNP is correlated with an

CC altered risk for myocardial infarction in the individual. The invention

CC further comprises: an isolated nucleic acid molecule comprising at least

CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in

CC the specification or its complement and encoding any one of the amino

CC acid sequences given in the specification; an isolated polypeptide

CC comprising an amino acid sequence given in the specification; an antibody

CC that specifically binds to the polypeptide or its antigen-binding

CC fragment; an amplified polynucleotide containing an SNP given in the

CC specification and which is between about 16 and 1000 nucleotides in

CC length; a kit for detecting an SNP in a nucleic acid, comprising the

CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a

CC nucleic acid molecule; a method of detecting a variant polypeptide; and a

CC method for identifying an agent useful in treating or preventing

CC myocardial infarction. The novel detection method has cardiac activity.

CC The nucleic acids of the invention may be used in gene therapy. The

CC method is useful in identifying an individual who has an increased or

CC decreased risk for developing myocardial infarction and for preparing a

CC composition for treating or preventing myocardial infarction. This

CC sequence represents the protein of a human myocardial infarction-

CC associated gene containing one or more SNPs of the invention. Note: This

CC sequence was not shown in the specification. The sequence has come from

CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 432 AA;

Query Match 94.9%; Score 56; DB 8; Length 432;

Best Local Similarity 91.7%; Pred. No. 0.022;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12

| | | | | : | | | | |

Db 208 FQGVLSVRVFV 219

RESULT 8

ADQ39357

ID ADQ39357 standard; protein; 432 AA.

XX ADQ39357;

XX ADQ39357;

XX ADQ39357;

DT 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

DE Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

OS Homo sapiens.

XX Homo sapiens.

XX WO2004058052-A2.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX Cargill M, Devlin JJ, Iakoubova O;

XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38529.
XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PS the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1020; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
XX CC has an altered risk for developing myocardial infarction. The method
XX CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX CC the nucleotide sequences given in the specification in the individual's
XX CC nucleic acids, where the presence of the SNP is correlated with an
XX CC altered risk for myocardial infarction in the individual. The invention
XX CC further comprises: an isolated nucleic acid molecule comprising at least
XX CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
XX CC the specification or its complement and encoding any one of the amino
XX CC acid sequences given in the specification; an isolated polypeptide
XX CC comprising an amino acid sequence given in the specification; an antibody
XX CC that specifically binds to the polypeptide or its antigen-binding
XX CC fragment; an amplified polynucleotide containing an SNP given in the
XX CC specification and which is between about 16 and 1000 nucleotides in
XX CC length; a kit for detecting an SNP in a nucleic acid, comprising the
XX CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
XX CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
XX CC method for identifying an agent useful in treating or preventing
XX CC myocardial infarction. The novel detection method has cardiant activity.
XX CC The nucleic acids of the invention may be used in gene therapy. The
XX CC method is useful in identifying an individual who has an increased or
XX CC decreased risk for developing myocardial infarction and for preparing a
XX CC composition for treating or preventing myocardial infarction. This
XX CC sequence represents the protein of a human myocardial infarction-
XX CC associated gene containing one or more SNPs of the invention. Note: This
XX CC sequence was not shown in the specification. The sequence has come from
XX CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;
Query Match 94.9%; Score 56; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. NO. 0.022;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FQGVLSQVRVFV 12
Db 208 FQGVLSQVRVFV 219
RESULT 9
AAU02916
ID AAU02916 standard; protein; 459 AA.
XX AC AAU02916;
XX DT 12-SEP-2001 (first entry)
XX DE Angiotensin converting enzyme (ACEV) splice variant protein #16.
XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX KW platelet-derived endothelial cell growth factor; cardiovascular disease;
XX KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX KW myocardial infarction; coronary arterial thrombosis; renal disease;
XX KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
XX KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
XX KW vascular disorder; asbestosis.

XX OS Homo sapiens.
XX FN WO200136632-A2.
XX PD 25-MAY-2001.
XX PF 17-NOV-2000; 2000WO-IL000766.
XX PR 17-NOV-1999; 99IL-00132978.
XX PR 10-DEC-1999; 99IL-00133455.
XX PA (COMP-) COMPUEN LTD.
XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX DR WPI; 2001-336004/35.
XX DR N-PSDB; AAS06016.
XX PT Novel alternative splicing variants e.g. variant of angiotensin
XX PT converting enzyme (ACEV), useful in identifying candidate compounds
XX PT capable of binding to the variant and to detect anti-variant antibodies.
XX PS Claim 4; Fig 16; 519pp; English.
XX CC The sequence represents an angiotensin converting enzyme splice variant
XX CC (ACEV) polypeptide. The polypeptides of the invention include variants of
XX CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
XX CC polypeptide receptor 2. The polypeptides and their associated nucleic
XX CC acids are useful for identification of variant sequences and detection of
XX CC candidate compounds capable of binding the molecules. The sequences of
XX CC the invention can be used in the treatment and diagnosis of various
XX CC disorders including cardiovascular diseases such as arteriosclerosis,
XX CC myocardial infarction and coronary arterial thrombosis, renal diseases
XX CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX CC immune disorders such as immune complex nephritis, multiple sclerosis,
XX CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
XX CC as asbestosis and vascular pathologies involving an endothelial
XX CC abnormality such as deep vein thrombosis
XX SQ Sequence 459 AA;
Query Match 94.9%; Score 56; DB 4; Length 459;
Best Local Similarity 91.7%; Pred. NO. 0.024;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FQGVLSQVRVFV 12
Db 208 FQGVLSQVRVFV 219
RESULT 10
AAB43602
ID AAB43602 standard; protein; 466 AA.
XX AC AAB43602;
XX DT 08-FEB-2001 (first entry)
XX DE Human cancer associated protein sequence SEQ ID NO:1047.
XX KW Human; cancer associated gene; cancer antigen; detection; cancer;
XX KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
XX KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
XX KW antiinflammatory; antichryoid; antiallergic; antibacterial; cardiac;
XX KW dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
XX KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX KW allergic reaction; graft versus host disease; organ rejection;
XX KW haemostatic; thrombolytic; cardiovascular disorder; infection;
XX KW neurological disease; drug screening.

```

XX OS Homo sapiens.
XX KW WO200055350-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005882.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX DR N-PSDB; AAC77811.
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX PS Claim 11; Page 1636-1638; 2352pp; English.
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
CC dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX SQ Sequence 466 AA;
Query Match 94.9%; Score 56; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.024;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FQGVLSQVRVFV 12
Db 261 FQGVLSQVRVFV 272
RESULT 11
AAU02915
ID AAU02915 standard; protein; 546 AA.
XX AC AAU02915;
XX DT 12-SEP-2001 (first entry)
XX DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX KW platelet-derived endothelial cell growth factor; cardiovascular disease;
XX KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX KW myocardial infarction; coronary arterial thrombosis; renal disease;
XX KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

```

```

KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX OS Homo sapiens.
XX FN WO200136632-A2.
XX PD 25-MAY-2001.
XX PF 17-NOV-2000; 2000WO-IL000766.
XX PR 17-NOV-1999; 99IL-00132978.
XX PR 10-DEC-1999; 99IL-00133455.
XX PA (COMP-) COMFUGEN LTD.
XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX DR N-PSDB; AAS06015.
XX PS Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
XX capable of binding to the variant and to detect anti-variant antibodies.
XX Claim 4; Fig 15; 519pp; English.
XX CC The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
XX polypeptide receptor 2. The polypeptides and their associated nucleic
XX acids are useful for identification of variant sequences and detection of
XX candidate compounds capable of binding the molecules. The sequences of
XX the invention can be used in the treatment and diagnosis of various
XX disorders including cardiovascular diseases such as arteriosclerosis,
XX myocardial infarction and coronary arterial thrombosis, renal diseases
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX immune disorders such as immune complex nephritis, multiple sclerosis,
XX cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
XX as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
XX SQ Sequence 546 AA;
Query Match 94.9%; Score 56; DB 4; Length 546;
Best Local Similarity 91.7%; Pred. No. 0.029;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FQGVLSQVRVFV 12
Db 208 FQGVLSQVRVFV 219
RESULT 12
ADN02474
ID ADN02474 standard; protein; 548 AA.
XX AC ADN02474;
XX DT 17-JUN-2004 (first entry)
XX DE TSF polypeptide.
XX KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX OS Homo sapiens.
XX FN CN1401387-A.
XX PD 12-MAR-2003.

```

XX 21-AUG-2002; 2002CN-00129408.
 XX
 XX 21-AUG-2002; 2002CN-00129408.
 XX
 XX (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
 XX
 XX Han Z, Liu P;
 XX
 XX WPI; 2003-469302/45.
 XX
 XX N-PSDB; ADN02475.
 XX
 XX Tumor suppressing polypeptide TSF and gene therapy vector composition.
 XX
 XX Claim 2; SEQ ID NO 1; 13pp; Chinese.
 XX
 XX The present invention relates to a novel recombinant adenovirus vector
 CC mediated anti-neoplastic composition is prepared through cloning the cDNA
 CC sequence from the human peripheral blood cell by specific primer and
 CC reverse transcription-polymerase chain reaction (RT-PCR) method for
 CC coding TSF polypeptide, construction in human embryonic kidney cell 293
 CC by AdEasy system, and packaging and expressing the recombinant adenovirus
 CC vector of TSF. It can suppress the growth and transfer of cancer. The
 CC present sequence represents the TSF polypeptide.
 XX
 XX Sequence 548 AA;
 SQ

Query Match 94.9%; Score 56; DB 7; Length 548;
 Best Local Similarity 91.7%; Pred. No. 0.029;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVFV 12
 |||||:|||||
 Db 208 FQGVLSQVRVFV 219

RESULT 13
 AAU02914
 ID AAU02914 standard; protein; 555 AA.
 XX
 XX AC AAU02914;
 XX
 XX DT 12-SEP-2001 (first entry)
 XX
 XX DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
 XX
 XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 XX OS Homo sapiens.
 XX
 XX PN W0200136632-A2.
 XX
 XX PD 25-MAY-2001.
 XX
 XX PF 17-NOV-2000; 2000WO-IL000766.
 XX
 XX PR 17-NOV-1999; 99IL-00132978.
 XX
 XX PR 10-DEC-1999; 99IL-00133455.
 XX
 XX PA (COMP-) COMPUEN LTD.
 XX
 XX PI Levine Z, David A, Azar I, Khooravi R, Bernstein J;
 XX
 XX WPI; 2001-336004/35.
 XX
 XX DR

DR N-PSDB; AAS06014.
 XX
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 XX Claim 4; Fig 14; 519pp; English.
 XX
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 XX Sequence 555 AA;
 SQ

Query Match 94.9%; Score 56; DB 4; Length 555;
 Best Local Similarity 91.7%; Pred. No. 0.03;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVFV 12
 |||||:|||||
 Db 208 FQGVLSQVRVFV 219

RESULT 14
 AAU02913
 ID AAU02913 standard; protein; 731 AA.
 XX
 XX AC AAU02913;
 XX
 XX DT 12-SEP-2001 (first entry)
 XX
 XX DE Angiotensin converting enzyme (ACEV) splice variant protein #13.
 XX
 XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 XX OS Homo sapiens.
 XX
 XX PN W0200136632-A2.
 XX
 XX PD 25-MAY-2001.
 XX
 XX PF 17-NOV-2000; 2000WO-IL000766.
 XX
 XX PR 17-NOV-1999; 99IL-00132978.
 XX
 XX PR 10-DEC-1999; 99IL-00133455.
 XX
 XX PA (COMP-) COMPUEN LTD.
 XX
 XX PI Levine Z, David A, Azar I, Khooravi R, Bernstein J;
 XX
 XX WPI; 2001-336004/35.
 XX
 XX DR

DR N-PSDB; AAS06013.
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 PS Claim 4; Fig 13; 519pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 731 AA;

Query Match 94.9%; Score 56; DB 4; Length 731;
 Best Local Similarity 91.7%; Pred. No. 0.04; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 1;

Qy 1 FQGVLSQSVRFVF 12
 |||||:||||
 Db 208 FQGVLSQSVRFVF 219

RESULT 15
 AAB00042
 ID AAB00042 standard; protein; 1152 AA.
 XX
 AC AAB00042;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Human thrombospondon-1 (TSP-1).
 XX
 KW TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
 KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
 KW diabetic retinopathy; corneal graft rejection; glaucoma.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Region 361..416
 FT /label= Type 1 repeat region
 FT Region 417..473
 FT /label= Type 1 repeat region
 FT Region 474..530
 FT /label= Type 1 repeat region

XX WO200044908-A2.

XX 03-AUG-2000.

XX 01-FEB-2000; 2000WO-US002482.

XX 01-FEB-1999; 99US-0118053P.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Lawler JW;

XX WPI; 2000-514823/46.

XX Nucleic acids encoding chimeric proteins such as cartilage oligomeric
 PT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
 PT inhibiting angiogenesis and treating diseases such as cancer.
 XX
 PS Disclosure; Fig 1; 40pp; English.

XX New nucleic acids are described which encode a protein comprising the
 CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
 CC the TGF (transforming growth factor)-beta activation region of human TSP-
 CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
 CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
 CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
 CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
 CC tumours in mice models. Thus the nucleic acids and proteins may be useful
 CC for treating angiogenesis related diseases such as cancer (by reducing
 CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
 CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
 CC for treating human immunodeficiency virus (HIV) infection. Anti-
 CC angiogenic therapy has little toxicity, does not require the therapeutic
 CC agent to enter tumour cells or cross the blood-brain barrier, controls
 CC tumour growth independently of growth of tumour cell heterogeneity, and
 CC does not induce drug resistance

XX SQ Sequence 1152 AA;

Query Match 94.9%; Score 56; DB 3; Length 1152;

Best Local Similarity 91.7%; Pred. No. 0.066;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQSVRFVF 12
 |||||:||||
 Db 190 FQGVLSQSVRFVF 201

Search completed: June 5, 2006, 22:24:58

Job time : 91.1379 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-21

Perfect score: 59

Sequence: 1 FQGVLSQVRVVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	94.9	229	2	thrombospondin 1 -
2	56	94.9	1170	1	thrombospondin 1 p
3	56	94.9	1170	2	thrombospondin 1 p
4	38	64.4	145	2	hypothetical prote
5	38	64.4	162	2	protein F1504.3 (i
6	38	64.4	212	2	probable imidazole
7	38	64.4	747	2	two component resp
8	38	64.4	783	2	probable transcrip
9	38	64.4	1172	1	thrombospondin 2 p
10	38	64.4	1172	2	thrombospondin 2 p
11	37	62.7	715	2	zona pellucida gly
12	36	61.0	111	2	probable membrane
13	36	61.0	151	2	tRNA-pseudouridine
14	36	61.0	993	2	hypothetical prote
15	36	61.0	1142	2	probable negative
16	36	61.0	1178	1	thrombospondin pre
17	36	61.0	1312	2	DNA-directed DNA p
18	36	61.0	1670	2	DNA-directed DNA p
19	35	59.3	82	2	hypothetical prote
20	35	59.3	465	2	probable membrane
21	35	59.3	681	2	hypothetical prote
22	35	59.3	819	2	hypothetical prote
23	34	57.6	212	2	hypothetical prote
24	34	57.6	265	2	hypothetical prote
25	34	57.6	348	2	p-aminobenzoate sy
26	34	57.6	355	2	G protein alpha ch
27	34	57.6	363	2	PTS system, fructo
28	34	57.6	364	2	PTS system, fructo
29	34	57.6	417	2	hypothetical prote

30	34	57.6	863	2	S41984
31	34	57.6	893	2	S36795
32	34	57.6	1702	2	S42459
33	33.5	56.8	827	2	A95877
34	33	55.9	106	2	C90261
35	33	55.9	167	2	A11167
36	33	55.9	307	2	D70384
37	33	55.9	311	2	F70184
38	33	55.9	372	2	T51082
39	33	55.9	397	2	G64703
40	33	55.9	406	2	G71816
41	33	55.9	459	2	D86669
42	33	55.9	467	2	D84938
43	33	55.9	469	2	T33595
44	33	55.9	471	2	S45068
45	33	55.9	505	2	E87021

ALIGNMENTS

RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C:Accession: S57957

R:Lafuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi

A:Reference number: S57955

A:Accession: S57957

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-229 <LAF>

A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:q899228; PDI

C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

Query Match 94.9%; Score 56; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0014;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12

Db 190 FQGVLSQVRVVF 201

RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)

C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A26155; A34274; A30140; A25812; A05172; A42927

R:Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple ce

A:Reference number: A26155; MUID:87057617; PMID:2430973

A:Accession: A26155

A:Molecule type: mRNA

A:Residues: 1-1170 <LAW>

A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:g37137; PIDN:C

A:Note: parts of this sequence, including the amino end of the mature protein, were deter

R:Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA se

A:Reference number: A34274; MUID:89291870; PMID:2544587

A:Accession: A34274

A:Molecule type: DNA

A:Residues: 1-166 <LAH>

A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835

R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I

J. Cell Biol. 108, 729-736, 1989

A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

A;Reference number: A30140; MUID:89139590; PMID:2918029
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-522','A','524-1170 <HEN>
A;Cross-references: UNIPARC:UPI0000038AB1; EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Kobayashi, S.; Eden-McCuthan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis o
A;Reference number: A25812; MUID:87157592; PMID:3030396
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-397 <KOB>
A;Cross-references: UNIPARC:UPI000016B0CA; GB:M25631; NID:G538353; PIDN:AAA36741.1; PID:
R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276; PMID:3461443
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-374','RC' <DIX>
A;Cross-references: UNIPARC:UPI000016B140; GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A;Reference number: A42927; MUID:92348511; PMID:1379247
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Cross-references: UNIPARC:UPI00001742C0
A;Note: Cys-992 is shown to have a free sulphydryl
C;Genetics:
A;Gene: GDB:THBS1; TSP1; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vc
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248,360,708,1067/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carboxydrate (Asn) (covalent) #status absent

Genomics 11, 587-600, 1991
A;Title: Characterization of the murine thrombospondin gene.
A;Reference number: A40558; MUID:92128941; PMID:1774063
A;Accession: A40558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1170 <LAW>
A;Cross-references: UNIPROT:P35441; UNIPARC:UPI0000028012; GB:M62449; GB:M62450; GB:M62451
R;GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469
P;Bornstein, P.; Alfli, D.; Devarayalu, S.; Framson, P.; Li, P.
J. Biol. Chem. 265, 16691-16698, 1990
A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of
A;Reference number: A37905; MUID:90375546; PMID:2398070
A;Accession: A37905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <BOR>
A;Cross-references: UNIPARC:UPI000016D076; GB:J05605; GB:J05606; NID:g201991; PIDN:AAA404
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: B42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1152,'P',1154-1170 <LAH>
A;Cross-references: UNIPARC:UPI0000177A96; GB:M87276
A;Note: sequence extracted from NCBI backbone (NCBIP:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 ar
A;Reference number: S68787; MUID:96234006; PMID:8654563
A;Accession: S68787
A;Molecule type: protein
A;Residues: 19-26,'X',28-37 <CHE>
A;Cross-references: UNIPARC:UPI0000177A97
C;Complex: homotrimer, disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor
C;Keywords: calcium binding; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carboxydrate (Asn) (covalent) #status predicted
Query Match 94.9%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.0079;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FQGVLSQVRVVF 12
|||||:|||||
Db 208 FQGVLSQVRVVF 219
RESULT 4
T20985
hypothetical protein F15D4.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20985
R;Smyle, R.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19354
A;Accession: T20985
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-145 <WIL>
A;Cross-references: UNIPROT:Q93511; UNIPARC:UPI000007BE33; EMBL:Z80344; PIDN:CAB02486.2;
A;Experimental source: clone F15D4
C;Genetics:

A:Gene: CESP:F15D4.3
A:Map position: 2
A:Introns: 21/3; 82/1

Query Match 64.4%; Score 38; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 3.6; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVF 12
Db 73 FMGVAQGLRYIF 84

RESULT 5
B88349
protein F15D4.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: B88349
R:anonymous, The C. elegans Sequencing Consortium.
C:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
Science 282, 2012-2018, 1998
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see website genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88349
A:Status: preliminary
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <STO>
A:Cross-references: UNIPROT:Q93511; UNIPARC:UPI0000179EEB; GB:chr_II; PIDN:CAB02486.1; F
C:Genetics:
A:Gene: F15D4.3
A:Map position: 2

Query Match 64.4%; Score 38; DB 2; Length 162;
Best Local Similarity 50.0%; Pred. No. 4.1; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVF 12
Db 90 FMGVAQGLRYIF 101

RESULT 6
D81929
probable imidazole-glycerol phosphate synthase amidotransferase component (EC 2.4.2.-) N
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: D81929
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81929
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <PAR>
A:Cross-references: UNIPROT:Q9JVH3; UNIPARC:UPI000012C805; GB:AL162754; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: hsh; NMA0840
C:Superfamily: amidotransferase hsh; trpG homology
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 64.4%; Score 38; DB 2; Length 212;
Best Local Similarity 58.3%; Pred. No. 5.4; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVF 12
Db 144 FQGIQDTRFYF 155

A;Accession: A47379
 A;Molecule type: mRNA
 A;Residues: 1-1172 <LA>
 A;Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046680; GB:L12350; NID:g307505; PIDN:R;LaBell, T.L.; Milewicz, D.J.; Distèche, C.M.; Byers, P.H.
 Genomics 12, 421-429, 1992
 A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression c
 A;Reference number: A42173; MUID:92217961; PMID:1559694
 A;Accession: A42173
 A;Molecule type: mRNA
 A;Residues: 560-1172 <LA2>
 A;Cross-references: UNIPARC:UPI00001742C1; GB:M81339
 A;Experimental source: fibroblast
 A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBI:P:95096)
 C;Genetics:
 A;Gene: GDB:THBS2; TSP2
 A;Cross-references: GDB:128789; OMIM:188061
 A;Map position: 6q27-6q27
 C;Complex: homotrimer, disulfide linked
 C;Function:
 A;Description: participates in cell migration and adhesion, and in platelet aggregation
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vc
 C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-1172/Product: thrombospondin 2 #status predicted <MAT>
 F;319-377/Domain: von Willebrand factor type C repeat homology <WVC>
 F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR2>
 F;553-588/Domain: EGF homology <EGF1>
 F;652-931/Domain: EGF homology <EGF>
 F;928-930/Region: cell attachment (R-G-D) motif
 F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predi
 F;167-226/Disulfide bonds: #status predicted
 F;266,270/Disulfide bonds: interchain #status predicted
 F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 64.4%; Score 38; DB 1; Length 1172;
 Best Local Similarity 58.3%; Pred. No. 33;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLQSVRFV 12
 Db 202 FRGLQNVHLVF 213

RESULT 10
 A42587
 thrombospondin 2 precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C;Accession: A42587; A39851
 R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
 J. Biol. Chem. 267, 3274-3281, 1992
 A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
 A;Reference number: A42587; MUID:92147683; PMID:1371115
 A;Accession: A42587
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: nucleic acid
 A;Residues: 1-1172 <LAH>
 A;Cross-references: UNIPROT:Q03350; UNIPARC:UPI000029847; GB:L07803; GB:M87275; NID:g34
 A;Note: sequence extracted from NCBI backbone (NCBI:P:81502)
 R;Bornstein, F.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
 J. Biol. Chem. 266, 12821-12824, 1991
 A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
 A;Reference number: A39851; MUID:91302287; PMID:1712771
 A;Accession: A39851
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-873 <BOR>
 A;Cross-references: UNIPARC:UPI0000160D77; GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vc
 C;Keywords: calcium binding; glycoprotein

F;319-377/Domain: von Willebrand factor type C repeat homology <WVC>
 F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F;553-588/Domain: EGF homology <EGF1>
 F;652-691/Domain: EGF homology <EGF>

Query Match 64.4%; Score 38; DB 2; Length 1172;
 Best Local Similarity 58.3%; Pred. No. 33;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLQSVRFV 12
 Db 202 FRGLQNVHLVF 213

RESULT 11
 S70397
 zona pellucida glycoprotein A - dog
 C;Species: Canis lupus familiaris (dog)
 C;Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
 C;Accession: S70397
 R;Harris, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.; Yurewicz, E.C.; Sacco, A.G.
 DNA Seq. 4, 361-393, 1994
 A;Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety of
 A;Reference number: S70396; MUID:95143578; PMID:7841460
 A;Accession: S70397
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-715 <HAR>
 A;Cross-references: UNIPROT:P47983; UNIPARC:UPI000013C45A; EMBL:U05779; NID:g458274; PID:
 C;Superfamily: sperm-binding glycoprotein ZP2; ZP domain homology
 F;368-628/Domain: ZP domain homology <ZPH>

Query Match 62.7%; Score 37; DB 2; Length 715;
 Best Local Similarity 80.0%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLQSVRF 10
 Db 277 FPGKLQSVRF 286

RESULT 12
 S52596
 Probable membrane protein YHR063w-a - yeast (Saccharomyces cerevisiae)
 C;Species: Saccharomyces cerevisiae
 C;Date: 05-May-1995 #sequence_revision 19-Oct-1995 #text_change 23-Mar-2001
 C;Accession: S52596
 R;Du, Z.
 A;Description: The sequence of S. cerevisiae cosmid 8025.
 A;Reference number: S46731
 A;Accession: S52596
 A;Molecule type: DNA
 A;Residues: 1-111 <DUZ>
 A;Cross-references: UNIPARC:UPI00001790B3; EMBL:U00061; GSPDB:GN00008; MIPS:YHR063w-a
 C;Genetics:
 A;Gene: MIPS:YHR063w-a
 A;Map position: 8R
 C;Superfamily: Saccharomyces probable membrane protein YHR063w-a
 C;Keywords: transmembrane protein
 F;49-65/Domain: transmembrane #status predicted <TMM>

Query Match 61.0%; Score 36; DB 2; Length 111;
 Best Local Similarity 50.0%; Pred. No. 7;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLQSVRFV 12
 Db 4 FQGLYSIMTVF 15

RESULT 13

C57253
tRNA-pseudouridine synthase (EC 5.4.99.-) - Acinetobacter calcoaceticus (fragment)
N:Alternate names: hypothetical protein lipB 5'-region
C:Species: Acinetobacter calcoaceticus
C>Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 05-Oct-2004
C:Accession: C57253
R:Kok, R.G.; van Thor, J.J.; Nugteren-Roodzant, I.M.; Vosman, B.; Hellingwerf, K.J.
J. Bacteriol. 177, 3295-3307, 1995
A:Title: Characterization of lipase-deficient mutants of Acinetobacter calcoaceticus BD4
A:Reference number: A57253; MUID:95286514; PMID:7768830
A:Accession: C57253
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-151 <KOK>
A:Cross-references: UNIPARC:UPI000017879A; GB:X80800
C:Superfamily: tRNA pseudouridine synthase B
C:Keywords: intramolecular transferase; isomerase; tRNA modification

Query Match 61.0%; Score 36; DB 2; Length 151;
Best Local Similarity 70.0%; Pred. No. 9.6;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12
||| |||:
Db 29 GVLQKRWLF 38

RESULT 14

T17230
hypothetical protein DKFZp434E066.1 - human
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T17230
R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17230
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-993 <KOE>
A:Cross-references: UNIPROT:Q9UES6; UNIPARC:UPI000016AC6D; EMBL:AL117432
A:Experimental source: adult testis; clone DKFZp434E066
C:Genetics:
A:Note: DKFZp434E066.1

Query Match 61.0%; Score 36; DB 2; Length 993;
Best Local Similarity 63.6%; Pred. No. 70;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12
||| |||:
Db 240 QHLLSLRFVF 250

RESULT 15

T39103
probable negative regulator of vesicle formation - fission yeast (Schizosaccharomyces po
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39103
R:Barrell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21828
A:Accession: T39103
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1142 <BAR>
A:Cross-references: UNIPROT:Q9UT41; UNIPARC:UPI000006A1AC; EMBL:AL121741; PIDN:CAB57332
A:Experimental source: strain 972h-; cosmid c824
C:Genetics:
A:Gene: SPDB:SPAC824.02
A:Map position: 1

C:Superfamily: Saccharomyces cerevisiae probable membrane protein YFL025C

Query Match 61.0%; Score 36; DB 2; Length 1142;
Best Local Similarity 58.3%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSQSVRFVF 12
||| |||:
Db 887 FTGVCQAVKFSF 898

Search completed: June 5, 2006, 22:44:58
Job time: 14.9655 secs

THIS PAGE BLANK (USPTO)

DE MGC64438 protein.
OS *Xenopus laevis* (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;

[1]

RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for *Xenopus* research: The NIH *Xenopus*
RT initiative";
RL Dev. Dyn. 225:384-391(2002).

[3]

RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----
CC EMBL; BC054970; AAH54970.1; -; mRNA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00090; TSP_1; 2.
CC Pfam; PF00093; VWC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 2.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PSS0092; TSP1; 2.
CC PROSITE; PS01208; VWFC_1; UNKNOWN_1.
CC PROSITE; PSS0184; VWFC_2; 1.
CC SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7EF51B CRC64;

Query Match 94.98; Score 56; DB 2; Length 496;
Best Local Similarity 91.74; Pred. No. 0.027; Indels 0; Gaps 0;
Matches 11; Conservative 1; Mismatches 0;

QY 1 FQGLQSVRFVF 12

|||||:|||||

Db 214 FQGLQSVRFVF 225

RESULT 3

TSPI_BOVIN
ID -TSPI_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSPI;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RX NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
RX MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Negata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
RT odontoblasts and predentin";
RL Biochim. Biophys. Acta 1382:17-22(1998).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGP-beta";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3. May play a role in dentinogenesis
CC and/or maintenance of dentin and dental pulp.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- TISSUE SPECIFICITY: Odontoblasts.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AB005287; BAA21115.1; -; mRNA.

EMBL; X87618; CAA60950.1; -; mRNA.

EMBL; X87619; CAA60951.1; -; mRNA.

PIR; S55501; S55501.

HSRP; P07996; 1LSL.

SMR; Q28178; 549-1169.

GlycoSuiteDB; Q28178; --

InterPro; IPR013320; ConA_like_subgrp.

InterPro; IPR006210; EGF.

InterPro; IPR000742; EGF 3.

InterPro; IPR001881; EGF_Ca_bd.

InterPro; IPR006209; EGF_like.

InterPro; IPR013032; EGF_like_reg.

InterPro; IPR003129; Laminin_G_TSP_N.

InterPro; IPR008085; TSP 1.

InterPro; IPR003367; Tsp_3.

InterPro; IPR008859; TSP_C.

InterPro; IPR001007; VWF_C.

Pfam; PF00008; EGF; 1.

Pfam; PF00090; TSP_1; 3.

Pfam; PF02412; TSP_3; 13.

DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWC; 1.
DR PROSITE; PS0184; VWC; 2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT REGION 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT CARBOHYD 1085 1085
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT DISULFID 519 531
FT DISULFID 551 562
FT DISULFID 556 572
FT DISULFID 575 586
FT DISULFID 592 608
FT DISULFID 599 617
FT DISULFID 620 644
FT DISULFID 650 663
FT DISULFID 657 676
FT DISULFID 678 689
FT DISULFID 705 713
FT DISULFID 718 738
FT DISULFID 754 774
FT DISULFID 777 797
FT DISULFID 813 833
FT DISULFID 836 856
FT DISULFID 874 894
FT DISULFID 910 930
FT DISULFID 946 1167
FT CONFLICT 805 805 S -> G (in Ref. 2).
SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 94.9%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.067;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVLSQVRVVF 12
DB 208 FQGVLSQVRVVF 219
RESULT 4
ID TSP1 HUMAN STANDARD; PRT; 1170 AA.
AC P07996; O15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
multiple calcium-binding sites and homologies with several different
proteins.";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
sites in the 3' untranslated region.";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Franson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
analysis of cDNA clones: homology to malarial circumsporozoite
proteins.";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/Genbank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;

RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.,
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT module.",
RN J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RN THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RT of the anti-angiogenic type 1 domains of human thrombospondin-1,";
RN Biochemistry 41:14329-14339(2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hydrazide chemistry, and mass spectrometry,";
RL J. Proteome Res. 4:2070-2080(2005).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIIb/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04665; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAA61237.1; ALT SEQ; mRNA.
DR EMBL; J04835; AAA61178.1; -; Genomic_DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1ZA4; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; Gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR013032; EGF-like_reg.
DR InterPro; IPR003129; Laminin_G_Tsp_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP-C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWFC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01184; VWFC_2; 1.
KW 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO_0000035842.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNac...).
FT N-linked (GlcNac...).
FT C-linked (Man).
FT /FTID=CAR_000205.
FT O-linked (Fuc...).
FT /FTID=CAR_000206.
FT C-linked (Man).
FT /FTID=CAR_000207.
FT C-linked (Man).
FT /FTID=CAR_000208.
FT O-linked (Fuc...).
FT /FTID=CAR_000209.
FT C-linked (Man).
FT /FTID=CAR_000210.
FT O-linked (Fuc...).
FT /FTID=CAR_000211.
FT N-linked (GlcNac...).
FT N-linked (GlcNac...).
FT Interchain (Probable).
FT Interchain (Probable).

Query Match 94.9%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.067;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12

Db 208 FQGVLSQVRVVF 219

RESULT 5

TSPI_MOUSE
 ID TSPI_MOUSE STANDARD; PRT; 1170 AA.
 AC P35441;
 DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
 DT 01-JUN-1994, sequence version 1.
 DT 07-MAR-2006, entry version 57.
 DE Thrombospondin-1 precursor.
 GN Name=Thbs1; Synonyms=Tspl;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92128941; PubMed=1774063;
 RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A.;
 RT "Characterization of the murine thrombospondin gene.";
 RL Genomics 11:587-600(1991).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92147883; PubMed=1371115;
 RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
 RA Dixit V.M.;
 RT "Characterization of mouse thrombospondin 2 sequence and expression
 during cell growth and development.";
 RL J. Biol. Chem. 267:3274-3281(1992).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 1-490.
 RX MEDLINE=90375546; PubMed=2398070;
 RA Bornstein P., Alfai D., Devarayalu S., Framson P., Li P.;
 RT "Characterization of the mouse thrombospondin gene and evaluation of
 the role of the first intron in human gene expression.";
 RL J. Biol. Chem. 265:16691-16698(1990).
 RN [4]
 RP PROTEIN SEQUENCE OF 19-37.
 RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
 RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
 RT "Expression and initial characterization of recombinant mouse
 thrombospondin 1 and thrombospondin 3.";
 RL FEBS Lett. 387:36-41(1996).
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
 cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
 laminin, type V collagen and integrins alpha-V/beta-1, alpha-
 V/beta-3 and alpha-IIB/beta-3.
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.
 CC -!- SIMILARITY: Belongs to the thrombospondin family.
 CC -!- SIMILARITY: Contains 3 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 VWFC domain.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 EMBL; M62470; AAA50611.1; -; Genomic DNA.
 DR EMBL; M62450; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62451; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62452; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62453; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62454; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62455; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62456; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62457; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62458; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62459; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62460; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; M62461; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62462; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62463; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62464; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62465; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62466; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62467; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62468; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M62469; AAA50611.1; JOINED; Genomic DNA.
 DR EMBL; M87276; AAA53063.1; -; mRNA.
 DR EMBL; J05606; AAA40431.1; -; Genomic DNA.
 DR PIR; A40558; A40558.
 DR HSP; P07996; ILSL.
 DR SMR; P35441; 549-1169.
 DR Ensembl; ENSMUSG00000040152; Mus musculus.
 DR MGI; MGI:98737; Thbs1.
 DR GO; GO:0005615; C:extracellular space; IDA.
 DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
 DR InterPro; IPR013320; ConA_like_subgrp.
 DR InterPro; IPR006210; EGF.
 DR InterPro; IPR00742; EGF 3.
 DR InterPro; IPR01881; EGF_Ca bd.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR013032; EGF-like_reg.
 DR InterPro; IPR003129; Laminin_G_TSP_N.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP 1.
 DR InterPro; IPR003367; tsp 3.
 DR InterPro; IPR008859; TSP_C.
 DR InterPro; IPR001007; VWFC_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 12.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; VWC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS01184; VWFC_2; 1.
 DR Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
 KW Glycoprotein; Heparin-binding; Repeat; Signal.
 KW SIGNAL
 FT CHAIN 1 18
 FT SIGNAL 19 1170
 FT Thrombospondin-1.
 FT /FTID=PRO_0000035843.
 FT TSP N-terminal.
 FT VWFC.
 FT TSP type-1 1.
 FT TSP type-1 2.
 FT TSP type-1 3.
 FT EGF-like 1.
 FT EGF-like 2; calcium-binding (Potential).
 FT EGF-like 3.
 FT TSP type-3 1.
 FT TSP type-3 2.
 FT TSP type-3 3.
 FT TSP type-3 4.
 FT TSP type-3 5.
 FT TSP type-3 6.
 FT TSP type-3 7.
 FT TSP C-terminal.
 FT Heparin-binding (Potential).
 FT Cell attachment site (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 248 248
 FT CARBOHYD 260 260
 FT CARBOHYD 708 708

```

FT CARBOHYD 1067 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 Interchain (Probable).
FT DISULFID 274 Interchain (Probable).
FT DISULFID 391 By similarity.
FT DISULFID 395 By similarity.
FT DISULFID 406 By similarity.
FT DISULFID 447 By similarity.
FT DISULFID 451 By similarity.
FT DISULFID 462 By similarity.
FT DISULFID 504 By similarity.
FT DISULFID 508 By similarity.
FT DISULFID 519 By similarity.
FT DISULFID 551 By similarity.
FT DISULFID 556 By similarity.
FT DISULFID 575 By similarity.
FT DISULFID 592 By similarity.
FT DISULFID 620 By similarity.
FT DISULFID 644 By similarity.
FT DISULFID 650 By similarity.
FT DISULFID 657 By similarity.
FT DISULFID 678 By similarity.
FT DISULFID 705 By similarity.
FT DISULFID 718 By similarity.
FT DISULFID 754 By similarity.
FT DISULFID 777 By similarity.
FT DISULFID 813 By similarity.
FT DISULFID 836 By similarity.
FT DISULFID 874 By similarity.
FT DISULFID 910 By similarity.
FT DISULFID 946 By similarity.
FT CONFLICT 1025 F -> L (in Ref. 2).
SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 94.9%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. NO. 0.067;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGLVQSVRFVF 12
Db 208 FQGLVQNVRFVF 219

RESULT 6
Q3TR40_MOUSE PRELIMINARY; PRT; 1170 AA.
AC Q3TR40;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 5.
DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DE clone:A530055N06 product:thrombospondin 1, full insert sequence.
GN Name=thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,

```

```

RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Heminger P., Girgeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnietz L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563 (2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566 (2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```


DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSEN; 1.
DR SMART; SM00214; VWC; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1; UNKNOWN_1.
DR PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 94.9%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.067;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVVF 12
|||||:|||||
Db 208 FQGVLSVRVVF 219

RESULT 8

ID Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwald J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted under the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; BC050917; AAH50917.1; -; mRNA.
DR HSSP; P07996; 1LSL.
DR SMR; Q80YQ1; 835-1170.
DR Ensembl; ENSMUSG0000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSEN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1; UNKNOWN_1.
DR PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 94.9%; Score 56; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. No. 0.067;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVVF 12
|||||:|||||
Db 208 FQGVLSVRVVF 219

RESULT 9

Q8CGB2 MOUSE
ID Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-
length enriched library, clone:G930018021 product:thrombospondin 1,
full insert sequence).
DE Name=Thbs1;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor
arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwald J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J.J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."; [2]
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAINS-CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor
RC arose spontaneously;
RA Strausberg R.; Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141072; DOI=10.1126/science.11120142;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Adinolfi V., Allen J.E.,
RA Ambesi-Impombato A., Aweiller N., Aturaliya R.N., Bailey T.L.,
RA Banal K.P., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnicki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kadowaki A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sander A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu Y., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=22354683; PubMed=12466951; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake E., Dragoti T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongsaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maitais L., Marchionni L., McKensie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sander A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wnshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willing L.,
RA Wnshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [9]


```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 602 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 94.9%; Score 56; DB 1; Length 1173;
Best Local Similarity 91.7%; Pred. No. 0.067;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLSQVRFVF 12
DB 211 FQGVLSQVRFVF 222

RESULT 11
Q59E99 HUMAN
ID Q59E99_HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AB209912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939; 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR008084; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
```

```
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_C_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 988BB16B57157B12 CRC64;

Query Match 94.9%; Score 56; DB 2; Length 1225;
Best Local Similarity 91.7%; Pred. No. 0.07;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLSQVRFVF 12
DB 263 FQGVLSQVRFVF 274

RESULT 12
Q5SPG5 BRARE
ID Q5SPG5_BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n\ (Fragment).
GN ORFName=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AL928866; CA120599.1; -; Genomic_DNA.
DR SMR; Q5SPG5; 751-804; 754-1089.
DR Ensembl; ENSDARG0000010785; Danio rerio.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR008084; TSP1.
DR InterPro; IPR008085; TSP_1.
```

```
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
DR Cell adhesion; EGF-like domain.
KW NON_TER 1
FT SEQUENCE 1090 AA; 120978 MW; 5A9320504A22DB936 CRC64;
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22DB936 CRC64;

Query Match 86.4%; Score 51; DB 2; Length 1090;
Best Local Similarity 83.3%; Pred. No. 0.64;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
Db 185 FMGVLSQVRVVF 196

RESULT 13
Q4S758_TETNG Q4S758_TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Chromosome 14 SCFA14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG0022976001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99983;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anchaud V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bismont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----

Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NoDerivs License
-----
CC EMBL; CAAE01014723; CAG03524.1; -; Genomic_DNA.
CC SMR; Q4S758; 811-1148.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR01881; EGF_Ca bd.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR00884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC Pfam; PF00093; VWC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS01186; EGF_2; UNKNOWN_1.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS01208; VWFC_1; 1.
CC PROSITE; PS0184; VWFC_2; 1.
CC Cell adhesion.
KW NON_TER 1
FT SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 86.4%; Score 51; DB 2; Length 1193;
Best Local Similarity 83.3%; Pred. No. 0.71;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
Db 187 FMGVLSQVRVVF 198

RESULT 14
Q5U903_PIG Q5U903_PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OC NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang X., Mauc G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AY773342; AAV38110.1; -; mRNA.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR001007; VWF_C.
```



```
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VMC_1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VMC_1; 1.
DR PROSITE; PS0184; VMC_2; 1.
FT NON_TER 249 249
SQ SEQUENCE 249 AA; 27560 MW; 465D664BE0329C32 CRC64;

Query Match      84.7%; Score 50; DB 2; Length 249;
Best Local Similarity 90.9%; Pred. No. 0.22;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVLSQVRVVF 12
   |||||:|||||
Db 1 QGVLSQVRVVF 11

RESULT 15
ID Q4RLR5_TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2006, sequence version 1.
DE 21-FEB-2006, entry version 8.
DE Chromosome 10 SCAP15019, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00032374001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dessat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volf J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quettier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAG01015019; CAG10667.1; -; Genomic_DNA.
DR SMR; Q4RLR5; 834-887, 837-1171.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005199; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
```

```
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR01881; EGF_Ca_bd.
DR InterPro; IPR013032; EGF like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR008884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR01007; VMC_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC_1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VMC_1; 1.
DR PROSITE; PS50184; VMC_2; 1.
KW Cell adhesion.
FT NON_TER 1171 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match      84.7%; Score 50; DB 2; Length 1171;
Best Local Similarity 83.3%; Pred. No. 1.1;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGVLSQVRVVF 12
   |||||:|||||
Db 207 FGVLSQVRVVF 218

Search completed: June 5, 2006, 22:42:38
Job time : 110.931 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-21

Perfect score: 59

Sequence: 1 FQGVLSVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5 COMB.pdp.*
 - 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6 COMB.pdp.*
 - 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7 COMB.pdp.*
 - 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H COMB.pdp.*
 - 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS COMB.pdp.*
 - 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE COMB.pdp.*
 - 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	94.9	825	2	US-09-949-002-482
2	56	94.9	831	2	US-09-939-853A-97
3	56	94.9	831	2	US-09-939-853A-98
4	56	94.9	1170	1	US-08-313-288B-20
5	56	94.9	1170	2	US-09-657-472-2
6	56	94.9	1170	2	US-09-949-002-350
7	38	64.4	175	2	US-09-252-991A-18826
8	38	64.4	1045	2	US-09-949-016-11112
9	38	64.4	1172	1	US-08-313-288B-19
10	38	64.4	1172	2	US-09-949-016-6333
11	37	62.7	158	2	US-09-248-796A-17510
12	37	62.7	287	2	US-09-258-634-5
13	37	62.7	715	1	US-08-484-932B-10
14	37	62.7	715	1	US-08-484-158B-10
15	37	62.7	715	1	US-08-484-596A-10
16	37	62.7	715	1	US-08-480-150A-10
17	37	62.7	715	2	US-08-458-731-10
18	37	62.7	715	2	US-08-149-223A-10
19	36	61.0	390	2	US-08-650-766-7
20	36	61.0	390	2	US-08-922-635-6
21	36	61.0	390	2	US-09-389-487-7
22	36	61.0	390	2	US-09-414-643-6
23	36	61.0	484	2	US-09-248-796A-16180
24	36	61.0	539	2	US-09-800-170-15
25	36	61.0	559	2	US-09-364-206-47
26	36	61.0	651	2	US-08-650-766-6

27	36	61.0	651	2	US-08-922-635-5	Sequence 5, Appli
28	36	61.0	651	2	US-09-389-487-6	Sequence 6, Appli
29	36	61.0	651	2	US-09-414-643-5	Sequence 5, Appli
30	36	61.0	1019	1	US-08-271-364A-7	Sequence 7, Appli
31	36	61.0	1019	1	US-08-222-715B-26	Sequence 26, Appli
32	36	61.0	1070	2	US-08-922-635-22	Sequence 22, Appli
33	36	61.0	1070	2	US-09-414-643-22	Sequence 22, Appli
34	36	61.0	1300	2	US-09-698-341-3	Sequence 3, Appli
35	36	61.0	1504	2	US-09-364-206-2	Sequence 2, Appli
36	35	59.3	99	2	US-09-270-767-61802	Sequence 61802, A
37	35	59.3	440	1	US-08-307-499-15	Sequence 15, Appli
38	35	59.3	440	2	US-09-299-268-15	Sequence 15, Appli
39	35	59.3	532	2	US-09-270-767-46234	Sequence 46234, A
40	35	59.3	944	2	US-09-134-000C-5578	Sequence 5578, Ap
41	34	57.6	169	2	US-09-134-000C-3511	Sequence 3511, Ap
42	34	57.6	731	2	US-09-270-767-42057	Sequence 42057, A
43	34	57.6	1006	2	US-09-710-279-154	Sequence 154, App
44	34	57.6	1022	1	US-08-271-364A-8	Sequence 8, Appli
45	34	57.6	1022	1	US-08-222-715B-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION

; FILE REFERENCE: AND USES THEREOF

; FILE REFERENCE: CL000790

; CURRENT APPLICATION NUMBER: US/09/949,002

; CURRENT FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/231,401

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 482

; LENGTH: 825

; TYPE: PRT

; ORGANISM: Human

US-09-949-002-482

Query Match 94.9%; Score 56; DB 2; Length 825;

Best Local Similarity 91.7%; Pred No. 0.0076; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12

|||||:|||||

Db 308 FQGVLSVRFVF 319

RESULT 2

US-09-939-853A-97

; Sequence 97, Application US/09939853A

; Patent No. 6989232

; GENERAL INFORMATION:

; APPLICANT: Burgess et al.

; TITLE OF INVENTION: NO. 6989232el Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-099

; CURRENT APPLICATION NUMBER: US/09/939,853A

; CURRENT FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: 60/228,191

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: 60/267,300

; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/269,961

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/277,337

; PRIOR FILING DATE: 2001-03-20

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          94.9%; Score 56; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.0076;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
      |||||:||||
Db      208 FQGVLSVRVFV 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          94.9%; Score 56; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.0076;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
      |||||:||||
Db      208 FQGVLSVRVFV 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          94.9%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
      |||||:||||
Db      208 FQGVLSVRVFV 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          94.9%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
      |||||:||||
Db      208 FQGVLSVRVFV 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-350

Query Match 94.9%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
|:|:|:|:|:
Db 208 FQGVLSQVRVVF 219

RESULT 7
US-09-252-991A-18826
; Sequence 18826, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18826
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18826

Query Match 64.4%; Score 38; DB 2; Length 175;
Best Local Similarity 63.6%; Pred. No. 6.2;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLSQVRVVF 12
|:|:|:|:|:
Db 41 QGVLSQVRVVF 51

RESULT 8
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11112

Query Match 64.4%; Score 38; DB 2; Length 1045;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
|:|:|:|:|:
Db 281 FRGLLQNVHLVF 292

RESULT 9
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match 64.4%; Score 38; DB 1; Length 1172;
Best Local Similarity 58.3%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
|:|:|:|:|:
Db 202 FRGLLQNVHLVF 213

RESULT 10
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6333

Query Match      64.4%; Score 38; DB 2; Length 1172;
Best Local Similarity 58.3%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 FQGVLSQVRVFV 12
      |||:||:|
Db      202 FRGLQNHLVF 213

RESULT 11
US-09-248-796A-17510
; Sequence 17510, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17510
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17510

Query Match      62.7%; Score 37; DB 2; Length 158;
Best Local Similarity 66.7%; Pred. No. 8.8;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 FQGVLSQVRVFV 12
      |||:||:|
Db      33 FQSVLSLRKVF 44

RESULT 12
US-09-258-634-5
; Sequence 5, Application US/09258634A
; Patent No. 6875855
; GENERAL INFORMATION:
; APPLICANT: Roberts, David
; APPLICANT: Yan, Sizhuang
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences Of
; Hemoglobin-Response Genes In Candida Albicans And The
; TITLE OF INVENTION: Use Of Reagents Derived From These Sequences In The
; Diagnosis Of disseminated Candida Albicans Infection
; FILE REFERENCE: 2026-4301
```

```
; CURRENT APPLICATION NUMBER: US/09/258,634A
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-258-634-5

Query Match      62.7%; Score 37; DB 2; Length 287;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 FQGVLSQVRVFV 12
      |||:||:|
Db      45 FQSVLSLRKVF 56

RESULT 13
US-08-484-993B-10
; Sequence 10, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 715 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-993B-10

Query Match      62.7%; Score 37; DB 1; Length 715;
Best Local Similarity 80.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 FQGVLSQVRVF 10
      ||| |||||
```

Db 277 FPGKLSVRF 286

RESULT 14
US-08-484-158B-10
; Sequence 10, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immunocontraception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-93
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 715 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-158B-10

Query Match 62.7%; Score 37; DB 1; Length 715;
Best Local Similarity 80.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FPGVLSVRF 10
DB 277 FPGKLSVRF 286

RESULT 15
US-08-484-596A-10
; Sequence 10, Application US/08484596A
; Patent No. 5981228
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.

; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/484,596A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 11-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 715 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-596A-10

Query Match 62.7%; Score 37; DB 1; Length 715;
Best Local Similarity 80.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FPGVLSVRF 10
DB 277 FPGKLSVRF 286

Search completed: June 5, 2006, 22:48:53
Job time : 24.8966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-21

Perfect score: 59

Sequence: 1 FQGVLSQVRVVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.Main.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	94.9	12	4	US-10-474-213-28
2	56	94.9	240	4	US-10-419-462-40
3	56	94.9	240	5	US-10-782-968-40
4	56	94.9	432	5	US-10-741-600-1020
5	56	94.9	432	5	US-10-741-600-1020
6	56	94.9	459	6	US-11-043-806-462
7	56	94.9	466	3	US-09-925-301-1047
8	56	94.9	555	6	US-11-043-806-454
9	56	94.9	578	6	US-11-043-806-456
10	56	94.9	685	6	US-11-043-806-452
11	56	94.9	804	6	US-11-043-806-453
12	56	94.9	828	6	US-11-043-806-455
13	56	94.9	831	3	US-09-939-853A-97
14	56	94.9	831	3	US-09-939-853A-98
15	56	94.9	855	6	US-11-043-806-461
16	56	94.9	1000	6	US-11-043-806-457
17	56	94.9	1105	6	US-11-043-806-458
18	56	94.9	1150	4	US-10-295-733-1
19	56	94.9	1152	3	US-09-919-603-1
20	56	94.9	1169	5	US-10-317-821B-7
21	56	94.9	1170	4	US-10-020-141-12
22	56	94.9	1170	4	US-10-017-721-2
23	56	94.9	1170	4	US-10-021-660-114
24	56	94.9	1170	4	US-10-008-093-2
25	56	94.9	1170	4	US-10-295-027-1170
26	56	94.9	1170	4	US-10-211-462-38
27	56	94.9	1170	4	US-10-231-956A-482

28	56	94.9	1170	4	US-10-419-462-38	Sequence 38, Appl
29	56	94.9	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	56	94.9	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	56	94.9	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	56	94.9	1170	5	US-10-782-968-38	Sequence 38, Appl
33	56	94.9	1170	5	US-10-849-989-44	Sequence 44, Appl
34	56	94.9	1170	5	US-10-631-467-548	Sequence 548, App
35	56	94.9	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	56	94.9	1170	5	US-10-831-997-2	Sequence 2, Appli
37	56	94.9	1170	5	US-10-995-561-594	Sequence 594, App
38	56	94.9	1170	5	US-10-995-561-595	Sequence 595, App
39	56	94.9	1170	5	US-10-995-561-596	Sequence 596, App
40	56	94.9	1170	6	US-11-037-713-51	Sequence 51, Appl
41	56	94.9	1170	6	US-11-046-644-28	Sequence 28, Appl
42	56	94.9	1170	6	US-11-046-456-28	Sequence 28, Appl
43	40	67.8	15	4	US-10-285-394-153	Sequence 153, App
44	39	66.1	226	5	US-10-467-657-2428	Sequence 2428, Ap
45	38	64.4	16	3	US-09-822-682-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Kruttsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/10535
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 94.9%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00074;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQVRVVF 12
Db 1 FQGVLSQVRVVF 12

RESULT 2

US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J.
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof in Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40

```
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match          94.9%; Score 56; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
Db 190 FQGVLSVRVFV 201

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match          94.9%; Score 56; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
Db 190 FQGVLSVRVFV 201

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match          94.9%; Score 56; DB 5; Length 432;
```

```
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match          94.9%; Score 56; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
Db 190 FQGVLSVRVFV 201

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match          94.9%; Score 56; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
Db 190 FQGVLSVRVFV 201

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match          94.9%; Score 56; DB 5; Length 432;
```

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 94.9%; Score 56; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.039;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
| | | | | : | | | | |
Db 261 FQGVLSVRVFV 272

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 94.9%; Score 56; DB 6; Length 555;
Best Local Similarity 91.7%; Pred. No. 0.047;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
| | | | | : | | | | |
Db 208 FQGVLSVRVFV 219

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 94.9%; Score 56; DB 6; Length 578;
Best Local Similarity 91.7%; Pred. No. 0.049;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
| | | | | : | | | | |
Db 208 FQGVLSVRVFV 219

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 94.9%; Score 56; DB 6; Length 685;
Best Local Similarity 91.7%; Pred. No. 0.059;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
| | | | | : | | | | |
Db 208 FQGVLSVRVFV 219

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 94.9%; Score 56; DB 6; Length 804;
Best Local Similarity 91.7%; Pred. No. 0.07;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12
| | | | | : | | | | |
Db 208 FQGVLSVRVFV 219

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match          94.9%; Score 56; DB 6; Length 828;
Best Local Similarity 91.7%; Pred. No. 0.072;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
Db      208 FQGVLSVRVFV 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          94.9%; Score 56; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.072;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
Db      208 FQGVLSVRVFV 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-09-939-853A-98

Query Match          94.9%; Score 56; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.072;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
Db      208 FQGVLSVRVFV 219

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847,1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match          94.9%; Score 56; DB 6; Length 855;
Best Local Similarity 91.7%; Pred. No. 0.074;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLSVRVFV 12
Db      208 FQGVLSVRVFV 219

Search completed: June 6, 2006, 00:00:09
Job time : 78.6207 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds

(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-21

Perfect score: 59

Sequence: 1 FQGVLSVRVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC Celerra SIDS3/ptodata/1/pubpa/US09_NEW_PUB.pdb:*
- 2: /EMC Celerra SIDS3/ptodata/1/pubpa/US06_NEW_PUB.pdb:*
- 3: /EMC Celerra SIDS3/ptodata/1/pubpa/US07_NEW_PUB.pdb:*
- 4: /EMC Celerra SIDS3/ptodata/1/pubpa/US08_NEW_PUB.pdb:*
- 5: /EMC Celerra SIDS3/ptodata/1/pubpa/PCT_NEW_PUB.pdb:*
- 6: /EMC Celerra SIDS3/ptodata/1/pubpa/US10_NEW_PUB.pdb:*
- 7: /EMC Celerra SIDS3/ptodata/1/pubpa/US11_NEW_PUB.pdb:*
- 8: /EMC Celerra SIDS3/ptodata/1/pubpa/US60_NEW_PUB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	61.0	1504	6	US-10-030-735-21
2	33	55.9	298	6	US-10-030-735-21
3	32	54.2	461	6	US-10-030-735-21
4	31	52.5	164	6	US-10-030-735-21
5	31	52.5	249	6	US-10-030-735-21
6	31	52.5	250	6	US-10-030-735-21
7	31	52.5	315	6	US-10-030-735-21
8	31	52.5	400	6	US-10-030-735-21
9	31	52.5	401	6	US-10-030-735-21
10	31	52.5	820	6	US-10-030-735-21
11	31	52.5	1043	6	US-10-030-735-21
12	31	52.5	3460	6	US-10-030-735-21
13	30	50.8	74	6	US-10-030-735-21
14	30	50.8	77	6	US-10-030-735-21
15	30	50.8	1333	6	US-10-030-735-21
16	29.5	50.0	101	6	US-10-030-735-21
17	29.5	50.0	161	6	US-10-030-735-21
18	29.5	50.0	189	6	US-10-030-735-21
19	29.5	50.0	258	6	US-10-030-735-21
20	29	49.2	15	7	US-11-247-376-5
21	29	49.2	139	7	US-11-293-697-3323
22	29	49.2	182	6	US-10-030-735-21
23	29	49.2	190	6	US-10-030-735-21
24	29	49.2	201	6	US-10-030-735-21
25	29	49.2	231	6	US-10-030-735-21

26	29	49.2	231	6	US-10-030-735-21
27	29	49.2	259	6	US-10-030-735-21
28	29	49.2	262	6	US-10-030-735-21
29	29	49.2	268	6	US-10-030-735-21
30	29	49.2	276	6	US-10-030-735-21
31	29	49.2	282	6	US-10-030-735-21
32	29	49.2	290	6	US-10-030-735-21
33	29	49.2	296	6	US-10-030-735-21
34	29	49.2	298	6	US-10-030-735-21
35	29	49.2	300	6	US-10-030-735-21
36	29	49.2	329	6	US-10-030-735-21
37	29	49.2	329	6	US-10-030-735-21
38	29	49.2	356	6	US-10-030-735-21
39	29	49.2	363	6	US-10-030-735-21
40	29	49.2	363	6	US-10-030-735-21
41	29	49.2	369	6	US-10-030-735-21
42	29	49.2	391	6	US-10-030-735-21
43	29	49.2	392	6	US-10-030-735-21
44	29	49.2	403	6	US-10-030-735-21
45	29	49.2	423	6	US-10-030-735-21

ALIGNMENTS

RESULT 1
US-10-505-928-662
; Sequence 662, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patent in 3.2
; SEQ ID NO 662
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-662

Query Match Similarity 61.0%; Score 36; DB 6; Length 1504;
Best Local Similarity 63.6%; Pred. NO. 28;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QGVLSVRVF 12
DB 751 QHILSLRFFV 761

RESULT 2
US-10-953-349-9347
; Sequence 9347, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579FUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 9347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9347

```
Query Match          55.9%; Score 33; DB 6; Length 298;
Best Local Similarity 40.0%; Pred. No. 18;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVQLSVRFV 11
   ::::|::|
Db 87 EGIQGVKFI 96

RESULT 3
US-10-505-928-282
; Sequence 282, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 282
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-282

Query Match          54.2%; Score 32; DB 6; Length 461;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVQLSVRFV 12
   ||::|::|
Db 375 FQGLLVSTIFCF 386

RESULT 4
US-10-953-349-1159
; Sequence 1159, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1159
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1159

Query Match          52.5%; Score 31; DB 6; Length 164;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVQLSVRFV 12
   |::|::|
Db 82 FEAVVDRVRLVF 93

RESULT 5
US-10-953-349-1158
; Sequence 1158, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
```

```
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1158
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1158

Query Match          52.5%; Score 31; DB 6; Length 249;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVQLSVRFV 12
   |::|::|
Db 167 FEAVVDRVRLVF 178

RESULT 6
US-10-953-349-1157
; Sequence 1157, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1157
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1157

Query Match          52.5%; Score 31; DB 6; Length 250;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVQLSVRFV 12
   |::|::|
Db 168 FEAVVDRVRLVF 179

RESULT 7
US-10-953-349-9204
; Sequence 9204, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9204
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9204

Query Match          52.5%; Score 31; DB 6; Length 315;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVQLSVRFV 12
```

```

; FILE REFERENCE: 259358USO
; CURRENT APPLICATION NUMBER: US/10/982,908
; CURRENT FILING DATE: 2004-11-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-908-26

Query Match          52.5%; Score 31; DB 6; Length 820;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1  FQGVLSRVFVF 12
      |||:|:| |
Db      123 FQGIQLINDFAY 134

RESULT 11
US-10-511-937-2452
; Sequence 2452, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2452
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2452

Query Match          52.5%; Score 31; DB 6; Length 1043;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      3  GVLQSRVFVF 12
      ||:|:|:|
Db      696 GILRTFKRIF 705

RESULT 12
US-10-505-928-104
; Sequence 104, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07

```

; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 104
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-104

Query Match 52.5%; Score 31; DB 6; Length 3460;
Best Local Similarity 54.5%; Pred. No. 6.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSRVFV 11
:|:|:|:|:
Db 2173 FEGQESDRFL 2183

RESULT 13

US-10-953-349-15283
; Sequence 15283, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15283
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15283

Query Match 50.8%; Score 30; DB 6; Length 74;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VLQSVRFV 12
:|:|:|:|:
Db 21 IVQGVRFAP 29

RESULT 14

US-10-953-349-15282
; Sequence 15282, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15282
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15282

Query Match 50.8%; Score 30; DB 6; Length 77;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VLQSVRFV 12
:|:|:|:|:
Db 24 IVQGVRFAP 32

RESULT 15
US-10-511-937-2992
; Sequence 2992, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2992
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2992

Query Match 50.8%; Score 30; DB 6; Length 1333;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSRVF 10
:|:|:|:|:
Db 327 FRGVLEQLRW 336

Search completed: June 6, 2006, 00:12:55
Job time : 3.82414 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 68.3534 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-22

Perfect score: 40

Sequence: 1 XGVLVNVRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- A_Geneseq_8.*
- 1: geneseqp1980s.*
 - 2: geneseqp1990s.*
 - 3: geneseqp2000s.*
 - 4: geneseqp2001s.*
 - 5: geneseqp2002s.*
 - 6: geneseqp2003as.*
 - 7: geneseqp2003bs.*
 - 8: geneseqp2004s.*
 - 9: geneseqp2005s.*
 - 10: geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	4 AAB35375	Aab35375 Alpha3bet
2	40	100.0	10	4 AAB35355	Aab35355 Alpha3bet
3	40	100.0	11	4 AAB35360	Aab35360 Alpha3bet
4	40	100.0	11	4 AAB35357	Aab35357 Alpha3bet
5	40	100.0	12	4 AAB35352	Aab35352 Alpha3bet
6	40	100.0	12	4 AAB35366	Aab35366 Alpha3bet
7	40	100.0	12	4 AAB35370	Aab35370 Alpha3bet
8	40	100.0	12	4 AAB35378	Aab35378 Alpha3bet
9	40	100.0	12	4 AAB35347	Aab35347 Alpha3bet
10	40	100.0	12	4 AAB35369	Aab35369 Alpha3bet
11	40	100.0	12	6 ABG72834	Abg72834 Thrombos
12	40	100.0	240	8 ADL70641	Adl70641 Human thr
13	40	100.0	432	8 ADQ39359	Adq39359 Human myo
14	40	100.0	432	8 ADQ39357	Adq39357 Human myo
15	40	100.0	459	4 AAU02916	Aau02916 Angiotens
16	40	100.0	466	3 AAB43602	Aab43602 Human can
17	40	100.0	546	4 AAU02915	Aau02915 Angiotens
18	40	100.0	548	7 ADN02474	Adn02474 TSF polyp
19	40	100.0	555	4 AAU02914	Aau02914 Angiotens
20	40	100.0	731	4 AAU02913	Aau02913 Angiotens
21	40	100.0	1152	3 AAB00042	Aab00042 Human thr
22	40	100.0	1152	5 AAU74771	Aau74771 Human thr
23	40	100.0	1152	5 ABB82285	Abb82285 Human thr

24	40	100.0	1170	4 AAB74450	Aab74450 Human var
25	40	100.0	1170	4 AAB90800	Aab90800 Human she
26	40	100.0	1170	5 AAE25030	Aae25030 Human thr
27	40	100.0	1170	5 AAU75315	Aau75315 Human thr
28	40	100.0	1170	6 ABP96780	Abp96780 Human COP
29	40	100.0	1170	6 ABU03474	Abu03474 Angiogene
30	40	100.0	1170	6 ABG74673	Abg74673 Human THB
31	40	100.0	1170	6 AAE36228	Aae36228 Human THB
32	40	100.0	1170	7 ABR62059	Abr62059 Human thr
33	40	100.0	1170	7 ADN39852	Adn39852 Cancer/an
34	40	100.0	1170	8 ADJ76124	Adj76124 Marker ge
35	40	100.0	1170	8 ADJ75296	Adj75296 Marker ge
36	40	100.0	1170	8 ADL70639	Adl70639 Human thr
37	40	100.0	1170	8 ADL35874	Adl35874 Human thr
38	40	100.0	1170	8 ADQ26070	Adq26070 Thrombos
39	40	100.0	1170	8 ADP54179	Adp54179 Human PRO
40	40	100.0	1170	8 ADQ39358	Adq39358 Human myo
41	40	100.0	1170	8 ADQ39356	Adq39356 Human myo
42	40	100.0	1170	8 ADQ39355	Adq39355 Human myo
43	40	100.0	1170	9 ADZ21688	Adz21688 Thrombos
44	40	100.0	1170	9 AEB87781	Aeb87781 Human thr
45	40	100.0	1170	9 AEB46751	Aeb46751 Human thr

ALIGNMENTS

RESULT 1

AAB35375

ID AAB35375 standard; peptide; 9 AA.

AC AAB35375;

DT 08-MAY-2001 (first entry)

DE Alpha3beta1 integrin binding peptide #40.

XX Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
XX diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
XX macular degeneration; psoriasis; cell adhesion; cell motility.
XX Synthetic.

XX WO2000105812-A2.

XX 25-JAN-2001.

XX 12-JUL-2000; 2000WO-US018986.

XX 15-JUL-1999; 99US-0144549P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roberts DD, Kruttsch HC;

XX WPI; 2001-182656/18.

XX New peptides that bind to or are recognized by alpha3-beta1 integrins,
XX useful for inhibiting cell adhesion to extracellular matrix, cell
XX motility and proliferation and for treating rheumatoid arthritis and
XX cancer.

XX Claim 4; Page 34; 84pp; English.

XX The present invention provides a number of peptides which bind to
XX alpha3beta1 integrins. They are useful in the modulation of cell adhesion
XX and motility, and in the treatment of cancer, diabetic retinopathy,
XX rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
XX and restenosis. The present sequence is an example of one of the peptides
XX of the invention

XX Sequence 9 AA;

Query Match 100.0%; Score 40; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.1e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
 Db 2 GVLQNVRF 9
 |||||

RESULT 2
 AAB35355
 ID AAB35355 standard; peptide; 10 AA.

XX AC AAB35355;
 XX DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #20.

XX KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
 XX KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 XX KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.
 XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX DR WPI; 2001-182656/18.

XX FT New peptides that bind to or are recognized by alpha3-betal integrins,
 PT useful for inhibiting cell adhesion to extracellular matrix, cell
 PT motility and proliferation and for treating rheumatoid arthritis and
 PT cancer.

XX PS Claim 4; Page 34; 84pp; English.

XX CC The present invention provides a number of peptides which bind to
 CC alpha3betal integrins. They are useful in the modulation of cell adhesion
 CC and motility, and in the treatment of cancer, diabetic retinopathy,
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
 CC and restenosis. The present sequence is an example of one of the peptides
 CC of the invention

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 40; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.058;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
 Db 3 GVLQNVRF 10
 |||||

RESULT 3
 AAB35360
 ID AAB35360 standard; peptide; 11 AA.

XX AC AAB35360;

XX DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #25.

XX KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
 KW KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 KW KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX DR WPI; 2001-182656/18.

XX FT New peptides that bind to or are recognized by alpha3-betal integrins,
 PT useful for inhibiting cell adhesion to extracellular matrix, cell
 PT motility and proliferation and for treating rheumatoid arthritis and
 PT cancer.

XX PS Claim 4; Page 34; 84pp; English.

XX CC The present invention provides a number of peptides which bind to
 CC alpha3betal integrins. They are useful in the modulation of cell adhesion
 CC and motility, and in the treatment of cancer, diabetic retinopathy,
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
 CC and restenosis. The present sequence is an example of one of the peptides
 CC of the invention

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 40; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
 Db 3 GVLQNVRF 10
 |||||

RESULT 4
 AAB35357

ID AAB35357 standard; peptide; 11 AA.

XX AC AAB35357;

XX DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #22.

XX KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
 KW KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 KW KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX WPI; 2001-182656/18.
XX New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX Claim 4; Page 34; 84pp; English.
XX The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX Sequence 11 AA;
SQ

Query Match 100.0%; Score 40; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GVLQNVRF 9
Db 2 GVLQNVRF 9
|||

RESULT 5
AAB35352
ID AAB35352 standard; peptide; 12 AA.
XX
AC AAB35352;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3betal integrin binding peptide #17.
XX
KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
FN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX Sequence 12 AA;
SQ

Query Match 100.0%; Score 40; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GVLQNVRF 9
Db 2 GVLQNVRF 9
|||

RESULT 6
AAB35366
ID AAB35366 standard; peptide; 12 AA.
XX
AC AAB35366;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3betal integrin binding peptide #31.
XX
KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
FN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX Sequence 12 AA;
SQ

Query Match 100.0%; Score 40; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GVLQNVRF 9
Db 3 GVLQNVRF 10
|||

RESULT 7
AAB35370
ID AAB35370 standard; peptide; 12 AA.
XX
AC AAB35370;
XX
DT 08-MAY-2001 (first entry)
XX
SQ Sequence 12 AA;
SQ

SQ Sequence 12 AA;
 Query Match 100.0%; Score 40; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.071;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
 Db 5 GVLQNVRF 12
 |||||

RESULT 10
 AAB35369
 ID AAB35369 standard; peptide; 12 AA.
 XX AC AAB35369;
 XX DT 08-MAY-2001 (first entry)
 XX DE Alpha3beta1 integrin binding peptide #34.
 XX KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
 KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 KW macular degeneration; psoriasis; cell adhesion; cell motility.
 XX OS Synthetic.
 XX PN WO200105812-A2.
 XX PD 25-JAN-2001.
 XX PF 12-JUL-2000; 2000WO-US018986.
 XX PR 15-JUL-1999; 99US-0144549P.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Roberts DD, Kruttsch HC;
 XX DR WPI; 2001-182656/18.
 PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
 PT useful for inhibiting cell adhesion to extracellular matrix, cell
 PT motility and proliferation and for treating rheumatoid arthritis and
 PT cancer.
 PS Claim 4; Page 34; 84pp; English.
 XX CC The present invention provides a number of peptides which bind to
 CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
 CC and motility, and in the treatment of cancer, diabetic retinopathy,
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
 CC and restenosis. The present sequence is an example of one of the peptides
 CC of the invention
 XX

SQ Sequence 12 AA;
 Query Match 100.0%; Score 40; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.071;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
 Db 3 GVLQNVRF 10
 |||||

RESULT 11
 ABG72834
 ID ABG72834 standard; peptide; 12 AA.
 XX AC ABG72834;
 XX DT 24-FEB-2003 (first entry)

XX Thrombospondin-1 sequence containing synthetic peptide.
 XX DE
 XX KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
 KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
 KW colon cancer; small-cell lung cancer; SCLC; melanoma.
 XX OS Synthetic.
 XX PN WO200281630-A2.
 XX PD 17-OCT-2002.
 XX PF 03-APR-2002; 2002WO-US010535.
 XX PR 06-APR-2001; 2001US-0281994P.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Roberts DD, Kruttsch HC;
 XX DR WPI; 2003-103329/09.
 XX PT A new diagnosis for cancer other than prostate cancer in a mammal useful
 PT to detect cancer including lung cancer, particularly small cell lung
 PT cancer and melanoma comprises detecting semenogelin in a sample.
 XX PS Example 1; Page 14; 32pp; English.
 XX CC The invention relates to diagnosing cancer other than prostate cancer in
 CC a male mammal, comprising assaying a test sample for increased level of
 CC semenogelin, or cancer in a female by assaying for the presence of
 CC semenogelin. Administering a semenogelin protein or polypeptide fragment
 CC or a semenogelin-specific antibody or active fragment, or a recombinant
 CC vector expressing the protein or antibody, is useful for inducing an
 CC immune response to a cancer in a mammal, where the cancer is not prostate
 CC cancer and semenogelin is a marker. The invention is used to diagnose
 CC cancer, particularly of epithelial origin such as lung cancer, papillary
 CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
 CC (SCLC), or a melanoma. The present sequence represents the amino acid
 CC sequence of the thrombospondin-1 sequence containing synthetic peptide
 CC which binds to alpha-3-beta-1 integrin
 XX

SQ Sequence 12 AA;
 Query Match 100.0%; Score 40; DB 6; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.071;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
 Db 3 GVLQNVRF 10
 |||||

RESULT 12
 ADL70641
 ID ADL70641 standard; protein; 240 AA.
 XX AC ADL70641;
 XX DT 20-MAY-2004 (first entry)
 XX DE Human thrombospondin-1 N-terminal domain.
 XX KW Human; thrombospondin-1; epitope; cancer; diagnosis.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Region 23..32
 FT /note= "Heparin binding region"
 FT Region 77..82
 FT /note= "Heparin binding region"

FT Region 151..164
PT /note= "Fibrinogen binding region"
XX
PN WO2004018995-A2.
XX
PD 04-MAR-2004.
XX
PF 20-AUG-2003; 2003WO-US026023.
XX
PR 23-AUG-2002; 2002US-0405494P.
PR 21-APR-2003; 2003US-00419462.
XX
PA (WILL/) WILLIAMS K J.
XX
XX Williams KJ;
PI
XX WPI; 2004-226901/21.
XX
DR New purified thrombospondin fragment extracted from a body fluid, useful
XX for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
PT or leukemia or as calibrators, indicators, immunogens and analytes.
PT
XX
PS Disclosure; SEQ ID NO 40; 76pp; English.
XX
XX The present sequence is that of the N-terminal domain of human
CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
CC clinical assays for cancer and for generation of antibodies and other
CC binding agents. A method that distinguishes TSP from a TSP fragment or
CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
CC or portion as a target for a binding molecule, e.g. an antibody, to
CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
CC epitope present in TSP but not in the fragment or portion to obtain a
CC quantitation of TSP only; and (3) using the difference between (1) and
CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
CC the TSP fragment or portion is performed in order to detect the presence,
CC or monitor the course, of a disease or condition selected from cancer,
CC renal failure, renal disease, atopic dermatitis, vasculitis, acute
CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
CC infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
CC nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition
CC associated with clotting, a condition associated with platelet
CC activation, a condition associated with intravascular platelet
CC activation, a condition associated with consumption of platelets, heparin
CC -induced thrombocytopenia, disseminated intravascular coagulation,
CC intravascular coagulation, extravascular coagulation, a condition
CC associated with endothelial activation, a condition associated with
CC production and/or release of thrombospondin and/or a thrombospondin
CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
CC reaction, an aspartame reaction, atopic dermatitis, eczema,
CC hypersensitivity, scleroderma, conditions associated with plugging of
CC vessels, a condition associated with a cryofibrinogen, a condition
CC associated with a cryoglobulin, and a condition associated with an anti-
CC cardiolin antibody. The cancer is selected from adenoma, cancer,
CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
CC the respiratory system, circulatory system, musculoskeletal system,
CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
CC biliary system, pancreas, head, neck, endocrine system, reproductive
CC system (male or female), genitourinary system, kidney, urinary tract,
CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
CC well-differentiated cancer or a moderately differentiated cancer.
XX
SQ Sequence 240 AA;

Query Match 100.0%; Score 40; DB 8; Length 240;

Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GVLQNVRF 9
Db 192 GVLQNVRF 199
RESULT 13
ADQ39359
ID ADQ39359 standard; protein; 432 AA.
XX AC ADQ39359;
XX DT 18-NOV-2004 (first entry)
XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1022.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human.
XX OS Homo sapiens.
XX FN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX N-PSDB; ADQ38531.
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 10; SEQ ID NO 1022; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNPs of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from

CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 432 AA;

Query Match 100.0%; Score 40; DB 8; Length 432;

Best Local Similarity 100.0%; Pred. No. 4.4; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0;

QY 2 GVLQNVRF 9

Db 210 GVLQNVRF 217

RESULT 14

ADQ39357

ID ADQ39357 standard; protein; 432 AA.

XX AC ADQ39357;

XX DT 18-NOV-2004 (first entry)

XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX KW cardiant; gene therapy; human.

XX OS Homo sapiens.

XX FN WO2004058052-A2.

XX PD 15-JUL-2004.

XX PF 22-DEC-2003; 2003WO-US040978.

XX PR 20-DEC-2002; 2002US-0434778P.

XX PR 10-MAR-2003; 2003US-0453135P.

XX PR 30-APR-2003; 2003US-0466412P.

XX PR 23-SEP-2003; 2003US-0504955P.

XX FA (APPL-) APPLERA CORP.

XX PI Cargill M, Devlin J, Iakoubova O;

XX DR WPI; 2004-533949/51.

XX DR N-PSDB; ADQ38529.

PT Identifying an individual who has an altered risk for developing myocardial infarction by detecting a single nucleotide polymorphism in the individual's nucleic acids.

XX Claim 10; SEQ ID NO 1020; 145pp; English.

XX The invention relates to a novel method for identifying an individual who has an altered risk for developing myocardial infarction. The method comprises detecting a single nucleotide polymorphism (SNP) in any one of the nucleotide sequences given in the specification in the individual's nucleic acids, where the presence of the SNP is correlated with an altered risk for myocardial infarction in the individual. The invention further comprises: an isolated nucleic acid molecule comprising at least 8 contiguous nucleotides where one of the nucleotides is an SNP given in the specification or its complement and encoding any one of the amino acid sequences given in the specification; an isolated polypeptide comprising an amino acid sequence given in the specification; an antibody that specifically binds to the polypeptide or its antigen-binding fragment; an amplified polynucleotide containing an SNP given in the specification and which is between about 16 and 1000 nucleotides in length; a kit for detecting an SNP in a nucleic acid, comprising the polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a nucleic acid molecule; a method of detecting a variant polypeptide; and a method for identifying an agent useful in treating or preventing myocardial infarction. The novel detection method has cardiant activity. The nucleic acids of the invention may be used in gene therapy. The method is useful in identifying an individual who has an increased or

CC decreased risk for developing myocardial infarction and for preparing a composition for treating or preventing myocardial infarction. This sequence represents the protein of a human myocardial infarction-associated gene containing one or more SNPs of the invention. Note: This sequence was not shown in the specification. The sequence has come from an electronic sequence listing downloaded from the WIPO website.

XX Sequence 432 AA;

Query Match 100.0%; Score 40; DB 8; Length 432;

Best Local Similarity 100.0%; Pred. No. 4.4; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0;

QY 2 GVLQNVRF 9

Db 210 GVLQNVRF 217

RESULT 15

AAU02916

ID AAU02916 standard; protein; 459 AA.

XX AC AAU02916;

XX DT 12-SEP-2001 (first entry)

XX DE Angiotensin converting enzyme (ACEV) splice variant protein #16.

XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;

XX KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;

XX KW platelet-derived endothelial cell growth factor; cardiovascular disease;

XX KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;

XX KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;

XX KW myocardial infarction; coronary arterial thrombosis; renal disease;

XX KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

XX KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;

XX KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

XX KW vascular disorder; asbestosis.

XX OS Homo sapiens.

XX PN WO200136632-A2.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-IL000766.

XX PR 17-NOV-1999; 99IL-00132978.

XX PR 10-DEC-1999; 99IL-00133455.

XX PA (COMP-) COMPUEN LTD.

XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX DR WPI; 2001-336004/35.

XX DR N-PSDB; AAS08016.

PT Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 4; Fig 16; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis,

CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 40; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
| | | | |
Db 210 GVLQNVRF 217

Search completed: June 5, 2006, 22:25:00
Job time : 69.3534 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 10.4741 Seconds

(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-22

Perfect score: 40

Sequence: 1 XGVLNQVRVF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	229	2 S57957	thrombospondin 1 -
2	40	100.0	1170	1 TSHUP1	thrombospondin 1 p
3	40	100.0	1170	2 A40558	thrombospondin 1 p
4	32	80.0	207	2 S53801	chitin synthase (E
5	32	80.0	780	2 T50315	hypothetical prote
6	32	80.0	889	2 JC6015	chitin synthase (E
7	32	80.0	1413	2 B82877	conserved hypothet
8	31	77.5	156	2 S60953	iron-sulfur cofact
9	31	77.5	186	2 B75421	probable pilin, ty
10	31	77.5	189	2 B45190	chitin synthase (E
11	31	77.5	195	2 H45189	chitin synthase (E
12	31	77.5	198	2 G45189	chitin synthase (E
13	31	77.5	198	2 A45190	chitin synthase (E
14	31	77.5	247	1 A64590	probable 3-oxoacyl
15	31	77.5	247	1 B71923	3-oxoacyl-[acyl-ca
16	31	77.5	300	2 D81399	malate dehydrogena
17	31	77.5	308	2 JC5468	leukocidin chain 1
18	31	77.5	311	2 C89968	leukotoxin Lute (i
19	31	77.5	387	1 TVECG	phosphoglycerate k
20	31	77.5	387	2 AD0875	phosphoglycerate k
21	31	77.5	387	2 E91103	phosphoglycerate k
22	31	77.5	387	2 H85948	phosphoglycerate k
23	31	77.5	387	2 AB0113	phosphoglycerate k
24	31	77.5	392	2 F82317	phosphoglycerate k
25	31	77.5	394	1 KIBSGM	phosphoglycerate k
26	31	77.5	394	2 C69675	phosphoglycerate k
27	31	77.5	406	2 G02022	tryptophan oxygena
28	31	77.5	446	2 G82299	phosphoglucumutase
29	31	77.5	467	2 DB4938	H+-transporting tw

30	31	77.5	747	2 AB2929	two component resp
31	31	77.5	783	2 A98352	probable transcrip
32	31	77.5	911	2 JC6016	chitin synthase (E
33	31	77.5	916	2 JC2315	chitin synthase (E
34	31	77.5	1114	2 T49517	p63 related protei
35	30	75.0	54	2 S35697	leukocidin chain F
36	30	75.0	70	2 F64066	probable outer mem
37	30	75.0	102	2 G84013	hypothetical prote
38	30	75.0	192	2 S77023	hypothetical prote
39	30	75.0	286	2 C49238	gamma-hemolysin co
40	30	75.0	310	2 S68225	synergohymenotropl
41	30	75.0	312	2 T00160	leukocidin chain S
42	30	75.0	312	2 S32211	leukocidin chain S
43	30	75.0	315	2 A49234	leucocidin R S com
44	30	75.0	315	2 JN0626	leukocidin chain S
45	30	75.0	315	2 E90043	gamma-hemolysin co

ALIGNMENTS

RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C:Accession: S57957

R:Lafuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi

A:Reference number: S57955

A:Accession: S57957

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-229 <LAF>

A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:G899228; PIDN:C

C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

Query Match 100.0%; Score 40; DB 2; Length 229;

Best Local Similarity 100.0%; Pred. No. 0.22;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9

Db 192 GVLQNVRF 199

RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)

C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A26155; A34274; A30140; A25812; A05172; A42927

R:Lawler, J.; Hynes, R.O.

J. Cell Biol. 103: 1635-1648, 1986

A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple ce

A:Reference number: A26155; MUID:87057617; PMID:2430973

A:Accession: A26155

A:Molecule type: mRNA

A:Residues: 1-1170 <LAW>

A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:G37137; PIDN:C

A>Note: parts of this sequence, including the amino end of the mature protein, were deter

R:Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264: 11222-11227, 1989

A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA se

A:Reference number: A34274; MUID:89291870; PMID:2544587

A:Accession: A34274

A:Molecule type: DNA

A:Residues: 1-166 <LAH>

A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835

R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I.

J. Cell Biol. 108: 729-736, 1989

A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

A:Gene: chsC
A:Function:
A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
A:Superfamily: chitin synthase chsA
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 80.0%; Score 32; DB 2; Length 207;
Best Local Similarity 85.7%; Pred. No. 14; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 1;

QY 2 GVLQNVNR 8
DB 16 GVMQNVNR 22

RESULT 5
T50315
hypothetical protein SPBC1703.02 with ARID DNA-binding domain [imported] - fission yeast
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50315
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Galibert, F.
submitted to the EMBL Data Library, January 2000
A:Reference number: Z25061
A:Accession: T50315
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-780 <MCD>
A:Cross-references: UNIPROT:Q9P7W8; UNIPARC:UPI000006A13D; EMBL:AL136536; PIDN:CAB66446.
A:Experimental source: strain 972h(-); cosmid cl703
C:Genetics:
A:Gene: SPDB:SPBC1703.02
A:Map position: 2
A:Introns: 38/2

Query Match 80.0%; Score 32; DB 2; Length 780;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
DB 94 GILQNVVF 101

RESULT 6
JC6015
chitin synthase (EC 2.4.1.16) C - Aspergillus fumigatus
N:Alternate names: chitin-UDP N-acetylglucosaminyltransferase C; CHSC protein
C:Species: Aspergillus fumigatus
C:Date: 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 09-Jul-2004
C:Accession: JC6015
R:Meillado, E.; Aufaivre-Brown, A.; Gow, N.A.R.; Holden, D.W.
Mol. Microbiol. 20, 667-679, 1996
A:Title: The Aspergillus fumigatus chsC and chsG genes encode class III chitin synthases
A:Reference number: JC6015; MUID:96347138; PMID:8736545
A:Accession: JC6015
A:Molecule type: mRNA
A:Residues: 1-889 <MEU>
A:Cross-references: UNIPROT:Q92197; UNIPARC:UPI000012794E; EMBL:X94245; NID:g1197185; PI
A:Experimental source: strain 237
C:Comment: This enzyme belongs to class III zymogen-type enzymes, it is a target for anti
C:Genetics:
A:Gene: chsC
A:Introns: 28/3; 169/2; 289/3; 633/2; 844/3
C:Function:
A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
C:Superfamily: chitin synthase chsA
C:Keywords: glycoprotein; Glycosyltransferase; hexosyltransferase; zymogen
F:524,845/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 80.0%; Score 32; DB 2; Length 889;
Best Local Similarity 85.7%; Pred. No. 70; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVLQNVNR 8
DB 218 GVMQNVNR 224

RESULT 7
B82877
conserved hypothetical UU543 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82877
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini
A:Reference number: A82870
A:Accession: B82877
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-1413 <GLA>
A:Cross-references: UNIPARC:UPI000000C1CB2; GB:AE002153; GB:AF222894; NID:g6899544; PIDN:
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UU543
A:Genetic code: SGC3

Query Match 80.0%; Score 32; DB 2; Length 1413;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VLQNVRF 9
DB 820 ILQNVRF 826

RESULT 8
S60953
iron-sulfur cofactor synthesis protein nifU homolog YOR226c [similarity] - yeast (Sacchar
N:Alternate names: hypothetical protein O5076; hypothetical protein YOR50-16
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 31-Dec-2004
C:Accession: S60953; S67119; S71728
R:Galisson, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A:Reference number: S60938
A:Molecule type: DNA
A:Residues: 1-156 <GAL>
A:Cross-references: UNIPROT:Q12056; UNIPARC:UPI00000698D8; EMBL:X92441; NID:g1050762; PI
R:Boyer, J.; Fairhead, C.; Gaillon, L.; Galisson, F.; Michaux, G.; Thierry, A.; Dujon, B.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67104
A:Accession: S67119
A:Molecule type: DNA
A:Residues: 1-156 <BOY>
A:Cross-references: UNIPARC:UPI00000698D8; EMBL:Z75133; NID:g1420519; PIDN:CAA99445.1; PI
A:Experimental source: strain S288C
R:Galisson, F.; Dujon, B.
Yeast 12, 877-885, 1996
A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of
A:Reference number: S71713; MUID:96437977; PMID:8840505
A:Accession: S71728
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-156 <GAW>
A:Cross-references: UNIPARC:UPI00000698D8; EMBL:X92441; NID:g1050762; PIDN:CAA63189.1; PI
C:Genetics:
A:Gene: SGD:ISU2
A:Cross-references: SGD:S0005752
A:Map position: 15R
C:Superfamily: iron-sulfur cluster assembly protein, iscU type; nitrogen fixation protein
C:Keywords: metalloprotein

F;60-113/Domain: nitrogen fixation protein homology <NFH>
F;61,88/Binding site: iron-sulfur clusters (Cys) (covalent) #status predicted

Query Match 77.5%; Score 31; DB 2; Length 156;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
|:|:|:|:
Db 76 GIENVRF 83

RESULT 9
B75421
C;Species: Deinococcus radiodurans (strain R1)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: B75421
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-186 <WHI>
A;Cross-references: UNIPROT:Q9RUZ7; UNIPARC:UPI00000D3E03; GB:AE001971; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1233
A;Map position: 1

Query Match 77.5%; Score 31; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 8
|:|:|:|:
Db 61 GVLNVRF 67

RESULT 10
B45190
C;Species: Exophiala dermatitidis (fragment)
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 03-Nov-2000
C;Accession: B45190
R;Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszló, P.J.; Robbins, P.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
A;Title: Classification of fungal chitin synthases.
A;Reference number: A38192; MUID:92115692; PMID:1731323
A;Accession: B45190
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-189 <BOW>
A;Cross-references: UNIPARC:UPI00001753F5
A;Note: sequence extracted from NCBI backbone (NCBIP:75855); the cited accession number,
A;Note: the source is designated as Wangiella dermatitidis
C;Superfamily: chitin synthase chsA
C;Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 77.5%; Score 31; DB 2; Length 189;
Best Local Similarity 71.4%; Pred. No. 21;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 8
|:|:|:|:
Db 10 GVMQNVRF 16

RESULT 11
A45190
C;Species: Rhinocladia atrovirens (fragment)
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 03-Nov-2000
C;Accession: A45190
R;Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszló, P.J.; Robbins, P.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
A;Title: Classification of fungal chitin synthases.
A;Reference number: A38192; MUID:92115692; PMID:1731323
A;Accession: A45190
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-198 <BOW>
A;Cross-references: UNIPARC:UPI00001753F3; GB:M82955; NID:g169826; PIDN:AAA33920.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIP:75854)
C;Superfamily: chitin synthase chsA
C;Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

H45189
C;Species: Ajellomyces capsulata - Ajellomyces capsulata (fragment)
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 03-Nov-2000
C;Accession: H45189
R;Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszló, P.J.; Robbins, P.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
A;Title: Classification of fungal chitin synthases.
A;Reference number: A38192; MUID:92115692; PMID:1731323
A;Accession: H45189
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-195 <BOW>
A;Cross-references: UNIPARC:UPI00001753BF; GB:M82948; NID:g168248; PIDN:AAA33381.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIP:75852)
C;Superfamily: chitin synthase chsA
C;Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 77.5%; Score 31; DB 2; Length 195;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 8
|:|:|:|:
Db 10 GVMQNVRF 16

RESULT 12
G45189
C;Species: Exophiala jeanselmei (fragment)
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 03-Nov-2000
C;Accession: G45189
R;Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszló, P.J.; Robbins, P.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
A;Title: Classification of fungal chitin synthases.
A;Reference number: A38192; MUID:92115692; PMID:1731323
A;Accession: G45189
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-198 <BOW>
A;Cross-references: UNIPARC:UPI00001753F4; GB:M82946; NID:g168142
A;Note: sequence extracted from NCBI backbone (NCBIP:75851); the sequence shown is from
C;Superfamily: chitin synthase chsA
C;Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 77.5%; Score 31; DB 2; Length 198;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 8
|:|:|:|:
Db 10 GVMQNVRF 16

RESULT 13
A45190
C;Species: Rhinocladia atrovirens (fragment)
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 03-Nov-2000
C;Accession: A45190
R;Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszló, P.J.; Robbins, P.W.
Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
A;Title: Classification of fungal chitin synthases.
A;Reference number: A38192; MUID:92115692; PMID:1731323
A;Accession: A45190
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-198 <BOW>
A;Cross-references: UNIPARC:UPI00001753F3; GB:M82955; NID:g169826; PIDN:AAA33920.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIP:75854)
C;Superfamily: chitin synthase chsA
C;Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Search completed: June 5, 2006, 22:45:04
Job time : 11.4741 secs

```

Query Match      77.5%; Score 31; DB 2; Length 198;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 GVLQNVRF 8
DB      10 GVMQNI 16

RESULT 14
A64590
probable 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Helicobacter pylori
C:Species: Helicobacter pylori
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: A64590
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: A64590
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-247 <TOM>
A:Cross-references: UNIPROT:O25286; UNIPARC:UPI00000D3178; GB:AE000570; GB:AE000511; NID
C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: fatty acid biosynthesis; NAD; oxidoreductase
F:6-186/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match      77.5%; Score 31; DB 1; Length 247;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVLQNVRF 9
DB      173 GALNRNRF 180

RESULT 15
B71923
3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Helicobacter pylori (strain
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 05-Oct-2004
C:Accession: B71923
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: B71923
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <ARN>
A:Cross-references: UNIPROT:Q9ZLS0; UNIPARC:UPI00000D3648; GB:AE001484; GB:AE001439; NID
A:Experimental source: strain J99
C:Genetics:
A:Gene: fabG
C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: oxidoreductase
F:6-186/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match      77.5%; Score 31; DB 2; Length 247;
Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVLQNVRF 9
DB      173 GALNRNRF 180

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 81.6983 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-22

Perfect score: 40

Sequence: 1 XGVQNVRF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	229	Q28194_BOVIN	Q28194 bos taurus
2	40	100.0	249	Q5U903_PIG	Q5U903 sus scrofa
3	40	100.0	496	Q7SY84_XENLA	Q7SY84 xenopus lae
4	40	100.0	1090	Q5SPG5_BRARE	Q5SPG5 brachydanio
5	40	100.0	1170	TSP1_BOVIN	Q28178 bos taurus
6	40	100.0	1170	TSP1_HUMAN	P07996 homo sapien
7	40	100.0	1170	TSP1_MOUSE	P35441 mus musculus
8	40	100.0	1170	Q3TR40_MOUSE	Q3TR40 mus musculus
9	40	100.0	1170	Q71SA3_RAT	Q71SA3 rattus norv
10	40	100.0	1171	Q80YQ1_MOUSE	Q80YQ1 mus musculus
11	40	100.0	1171	Q8CGB2_MOUSE	Q8CGB2 mus musculus
12	40	100.0	1171	Q4RL85_TETNG	Q4RL85 tetraodon n
13	40	100.0	1173	TSP1_XENLA	P35448 xenopus lae
14	40	100.0	1193	Q4S758_TETNG	Q4S758 tetraodon n
15	40	100.0	1225	Q59E99_HUMAN	Q59E99 homo sapien
16	39	97.5	704	Q4P665_USTMA	Q4P665 ustilago ma
17	39	97.5	1549	Q3CJX9_THEET	Q3CJX9 thermoanaer
18	37	92.5	1034	Q4RQ74_TETNG	Q4RQ74 tetraodon n
19	36	90.0	1172	Q6FXf4_CANGA	Q6FXf4 candida gla
20	36	90.0	1457	Q5UUF8_PRRSV	Q5UUF8 porcine rep
21	36	90.0	1457	Q6PR43_PRRSV	Q6PR43 porcine rep
22	36	90.0	1457	Q6TC18_PRRSV	Q6TC18 porcine rep
23	36	90.0	1457	Q91F53_PRRSV	Q91F53 porcine rep
24	36	90.0	1457	Q9DY97_PRRSV	Q9DY97 porcine rep
25	36	90.0	1457	Q9WBQ4_PRRSV	Q9WBQ4 porcine rep
26	36	90.0	1460	Q8QW9_PRRSV	Q8QW9 porcine rep
27	36	90.0	1463	Q4TUW0_PRRSV	Q4TUW0 porcine rep
28	36	90.0	1463	Q6QDRO_PRRSV	Q6QDRO porcine rep
29	36	90.0	1463	Q6SJE7_PRRSV	Q6SJE7 prsv hnl
30	36	90.0	1463	Q7TF56_PRRSV	Q7TF56 prsv hb-2(
31	36	90.0	1463	Q99AV5_PRRSV	Q99AV5 porcine rep

32	36	90.0	1463	2	Q99BU5_PRRSV	Q99BU5 porcine rep
33	36	90.0	1463	2	Q9EM9_PRRSV	Q9EM9 porcine rep
34	36	90.0	1463	2	Q9ENK5_PRRSV	Q9ENK5 porcine rep
35	36	90.0	3956	2	Q9DLN9_PRRSV	Q9DLN9 porcine rep
36	36	90.0	3960	1	RPOA_PRRSR	Q9WJb2 p replicase
37	36	90.0	3960	2	Q9DLN8_PRRSV	Q9WJb2 porcine rep
38	36	90.0	3960	2	Q9DLPO_PRRSV	Q9dlp0 porcine rep
39	36	90.0	3960	2	Q9DLPI_PRRSV	Q9dlp1 porcine rep
40	36	90.0	3963	1	RPOA_PRRSB	Q8B912 p replicase
41	36	90.0	3966	1	RPOA_PRRS1	Q9YN02 p replicase
42	35	87.5	713	2	Q3F1U8_9BURK	Q3f1u8 burkholderi
43	35	87.5	713	2	Q44XL2_9BURK	Q44xl2 burkholderi
44	35	87.5	713	2	Q4LLM8_9BURK	Q4llm8 burkholderi
45	35	87.5	1034	2	Q4DWH6_TRYCR	Q4Dwh6 trypanosoma

ALIGNMENTS

RESULT 1
Q28194_BOVIN
ID Q28194_BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96331130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCPI19>3.3.CO;2-0;
RA Lefeuvre B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR SMART; IPR003129; Laminin_G_TSP_N.
FT SMART: SM00210; TSPN; 1.
FT NON_TER 1 229
FT NON_TER 229 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9BECB4E8669C CRC64;

Query Match 100.0%; Score 40; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 192 GVLQNVRF 199
|||||||
192 GVLQNVRF 199

RESULT 2
Q5U903_PIG
ID Q5U903_PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.

DE Thrombospondin 1 (Fragment).
GN Name=Thb1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang K., Mauco G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

DR EMBL; AY773342; AAV38110.1; -; mRNA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
FT NON TER 1
FT NON TER 249 249
SQ SEQUENCE 249 AA; 27560 MW; 465D664B80329C32 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.4; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
| | | | | | | |
Db 2 GVLQNVRF 9

RESULT 3
Q7SY84 XENLA PRELIMINARY; PRT; 496 AA.
ID Q7SY84 XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Tissue=Whole;
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleorn M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshikiyuki S., Carninci F., Frange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

DR EMBL; BC054970; AAH54970.1; -; mRNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE; PS50184; VWFC_2; 1.
DR PROSITE; PS50184; VWFC_2; 1.
SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7EF51B CRC64;

Query Match 100.0%; Score 40; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
| | | | | | | |
Db 216 GVLQNVRF 223

RESULT 4
Q5SPG5 BRARE PRELIMINARY; PRT; 1090 AA.
ID Q5SPG5 BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.n\ (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AL928866; CAI20599.1; -; Genomic DNA.
DR SMR; Q5SPG5; 751-804, 754-1089.
DR Ensembl; ENSDARG0000010785; Danio rerio.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; C:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR002048; EF hand_Ca_bd.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF 3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF HAND 1; UNKNOWN 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_C; 1.
DR Cell adhesion; EGF-like domain.
FT NON TER 1
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1090;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
| | | | |
Db 187 GVLQNVRF 194

RESULT 5
ID TSP1_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
RX MEDLINE=9817373; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Nagata T., Teurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;

RT "cDNA cloning of bovine thrombospondin 1 and its expression in
ontoblasts and preterminal";
RL Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
TGF-beta.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
laminin, type V collagen and integrins alpha-V/beta-1, alpha-
V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis
and/or maintenance of dentin and dental pulp.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- TISSUE SPECIFICITY: Odontoblasts.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWF domain.

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AB005287; BAA21115.1; -; mRNA.
EMBL; X87618; CAA60950.1; -; mRNA.
EMBL; X87619; CAA60951.1; -; mRNA.
DR PIR; S55501; S55501.
DR HSP; P07996; ILSL.
DR SMR; Q28178; 549-1169.
DR GlycoSuiteDB; Q28178; -.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR03129; Laminin_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 13.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF 1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_C; 1.
DR Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
By similarity.
FT CHAIN 19 1170
Thrombospondin-1
/FTID=PRO_0000035841.
FT DOMAIN 24 221
TSP N-terminal.
FT DOMAIN 316 373
VWFC.
FT DOMAIN 379 429
TSP type-1 1.
FT DOMAIN 435 490
TSP type-1 2.
FT DOMAIN 492 547
TSP type-1 3.

```
FT DOMAIN 549 587 EGF-like 1.
FT DOMAIN 588 645 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 588 645 EGF-like 3.
FT DOMAIN 723 758 TSP type-3 1.
FT DOMAIN 759 781 TSP type-3 2.
FT DOMAIN 782 817 TSP type-3 3.
FT DOMAIN 818 840 TSP type-3 4.
FT DOMAIN 841 878 TSP type-3 5.
FT DOMAIN 879 914 TSP type-3 6.
FT DOMAIN 915 950 TSP type-3 7.
FT DOMAIN 951 1170 TSP C-terminal.
FT REGION 19 232 Cell attachment (Potential).
FT MOTIF 926 928 Cell attachment site (Potential).
FT CARBOHYD 248 248 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 360 360 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 708 708 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 1067 1067 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 1085 1085 N-linked (GlcNAc... ) (Potential).
FT DISULFID 270 270 Interchain (Probable).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 391 423 By similarity.
FT DISULFID 395 428 By similarity.
FT DISULFID 406 413 By similarity.
FT DISULFID 447 484 By similarity.
FT DISULFID 451 489 By similarity.
FT DISULFID 462 474 By similarity.
FT DISULFID 504 541 By similarity.
FT DISULFID 508 546 By similarity.
FT DISULFID 519 531 By similarity.
FT DISULFID 551 562 By similarity.
FT DISULFID 556 572 By similarity.
FT DISULFID 575 586 By similarity.
FT DISULFID 592 608 By similarity.
FT DISULFID 599 617 By similarity.
FT DISULFID 620 644 By similarity.
FT DISULFID 650 663 By similarity.
FT DISULFID 657 676 By similarity.
FT DISULFID 678 689 By similarity.
FT DISULFID 705 713 By similarity.
FT DISULFID 718 738 By similarity.
FT DISULFID 754 774 By similarity.
FT DISULFID 777 797 By similarity.
FT DISULFID 813 833 By similarity.
FT DISULFID 836 856 By similarity.
FT DISULFID 874 894 By similarity.
FT DISULFID 910 930 By similarity.
FT DISULFID 946 1167 By similarity.
FT CONFLICT 805 805 S -> G (in Ref. 2).
SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 100.0%; Score 40; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
    |||||
Db 210 GVLQNVRF 217

RESULT 6
TSPI1_HUMAN
ID TSPI1_HUMAN STANDARD; PRT; 1170 AA.
AC P07956; O15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DT 07-MAR-2006, entry version 78.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSF, TSPI;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
```

```
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L., Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan P., Framson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Glerman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441; THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;
RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J., Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1 module.";
RL J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments, of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E., Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry.";
RL J. Proteome Res. 4:2070-2080(2005).
CC -I- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibrinogen, fibronectin, laminin, type V collagen and integrins alpha-1/beta-1, alpha-
```


FT DISULFID 754 774 By similarity.
 FT DISULFID 777 797 By similarity.
 FT DISULFID 813 833 By similarity.
 FT DISULFID 836 856 By similarity.
 FT DISULFID 874 894 By similarity.
 FT DISULFID 910 930 By similarity.
 FT DISULFID 946 1167 By similarity.
 FT CONFLICT 1025 1025 F -> L (in Ref. 2).
 SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 1170;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
 |||||
 Db 210 GVLQNVRF 217

RESULT 8

Q3TR40 MOUSE
 ID Q3TR40_MOUSE PRELIMINARY; PRT; 1170 AA.
 AC Q3TR40;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
 DE clone:A530055N06 product:thrombospondin 1, full insert sequence.
 GN Name=Thbs1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Oyama R., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aurali R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Humnick L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jaki V.B., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Farzi S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Sempke C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
 RA Tannoja K., Tan S.L., Tang S., Tegner J., Taylor M.S., Tegner J.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,

RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brueic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 [3]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566(2005).
 [4]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kiyosawa H.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Gojobori T.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang N.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kageawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 [5]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=21085660; PubMed=1217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Fletcher C., Carninci P., de Bonaldo M.F.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,

```
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [6]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630 (2000).
RN [7]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsuoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ogawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:1757-1771 (2000).
RN [8]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK163092; BAE37190.1; -; mRNA.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPRO06210; EGF.
DR InterPro; IPRO00742; EGF_3.
DR InterPro; IPRO01881; EGF_Ca bd.
DR InterPro; IPRO06209; EGF like.
DR InterPro; IPRO13032; EGF like reg.
DR InterPro; IPRO03129; Laminin_G_TSP_N.
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO08085; TSP_1.
DR InterPro; IPRO03367; tsp_3.
DR InterPro; IPRO08859; TSP_C.
DR InterPro; IPRO01007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1170 AA; 129619 MW; 55BF0420D91B194 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 210 GVLQNVRF 217
|||||
|

RESULT 9
Q71SA3 RAT
ID Q71SA3 RAT PRELIMINARY; PRT; 1170 AA.
AC Q71SA3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 9.
DE Thrombospondin 1.
GN Name=tsp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX NCBI_TaxID=10116;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RA Iwabu A., Hirohata S., Kusachi S., Nakamura K., Murakami T.,
RA Ninomiya Y., Tsuji T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF109630; AAQ14549.1; -; mRNA.
DR SRR; Q71SA3; 834-1169.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPRO13320; ConA_like_subgrp.
DR InterPro; IPRO08210; EGF.
DR InterPro; IPRO00742; EGF_3.
DR InterPro; IPRO01881; EGF_Ca bd.
DR InterPro; IPRO06209; EGF like.
DR InterPro; IPRO13032; EGF like reg.
DR InterPro; IPRO03129; Laminin_G_TSP_N.
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO08085; TSP_1.
DR InterPro; IPRO03367; tsp_3.
DR InterPro; IPRO08859; TSP_C.
DR InterPro; IPRO01007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
```

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; WFC1; 1.
DR PROSITE; PS0184; WFC2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
|||||
Db 210 GVLQNVRF 217

RESULT 11
Q8CGB2 MOUSE
ID ID Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q8CGB2;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 21.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-length enriched library, clone:G930018021 product:thrombospondin 1, full insert sequence).
DE Name=Thbel;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RC Strausberg R.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL [3]
RN NUCLEOTIDE SEQUENCE.
RP

Db 210 GVLQNVRF 217

RESULT 10

Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.

ID Q80YQ1; Integrated into UniProtKB/TrEMBL.

AC Q80YQ1; sequence version 1.

DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.

DT 01-JUN-2003, sequence version 1.

DT 07-FEB-2006, entry version 17.

DE Thrombospondin 1.

DE Thrombospondin 1.

GN Name=Thbs1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridea; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=C57BL/6; TISSUE=Brain;

RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=C57BL/6; TISSUE=Brain;

RC Director MGC Project;

RX Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL; BC050917; AAH50917.1; -; mRNA.

DR HSSP; P07996; 1LSL.

DR SMR; Q80YQ1; 835-1170.

DR Ensembl; ENSMUSG00000040152; Mus musculus.

DR MGI; MGI:98737; Thbs1.

DR GO; GO:0005615; C:extracellular space; ICA.

DR GO; GO:0005615; C:extracellular space; ICA.

DR GO; GO:001525; P:negative regulation of angiogenesis; IDA.

DR InterPro; IPR013320; Cona_like_subgrp.

DR InterPro; IPR006210; EGF.

DR InterPro; IPR00742; EGF 3.

DR InterPro; IPR001881; EGF_Ca bd.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR013032; EGF_like_reg.

DR InterPro; IPR003129; Laminin_TSP_N.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP 1.

DR InterPro; IPR003367; tsp_3.

DR InterPro; IPR008859; TSP_C.

DR InterPro; IPR001007; VWF_C.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00090; TSP 1; 3.

DR

Query Match 100.0%; Score 40; DB 2; Length 1171;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9

DB 210 GVLQNVRF 217

RESULT 11

Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.

ID Q8CGB2; Integrated into UniProtKB/TrEMBL.

AC Q8CGB2;

DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.

DT 01-MAR-2003, sequence version 1.

DT 07-FEB-2006, entry version 22.

DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-

DE length enriched library, clone:G930018021 product:thrombospondin 1,

DE full insert sequence).

GN Name=Thbs1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridea; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor

RC arose spontaneously;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor

RC arose spontaneously;

RX Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

CC EMBL; BC050917; AAH50917.1; -; mRNA.

DR HSSP; P07996; 1LSL.

DR SMR; Q80YQ1; 835-1170.

DR Ensembl; ENSMUSG00000040152; Mus musculus.

DR MGI; MGI:98737; Thbs1.

DR GO; GO:0005615; C:extracellular space; ICA.

DR GO; GO:0005615; C:extracellular space; ICA.

DR GO; GO:001525; P:negative regulation of angiogenesis; IDA.

DR InterPro; IPR013320; Cona_like_subgrp.

DR InterPro; IPR006210; EGF.

DR InterPro; IPR00742; EGF 3.

DR InterPro; IPR001881; EGF_Ca bd.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR013032; EGF_like_reg.

DR InterPro; IPR003129; Laminin_TSP_N.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP 1.

DR InterPro; IPR003367; tsp_3.

DR InterPro; IPR008859; TSP_C.

DR InterPro; IPR001007; VWF_C.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00090; TSP 1; 3.

DR

- RC TISSUE=Mammary gland;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44 (1999).
RN [4]
- RC TISSUE=Mammary gland;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodius R., Shimokawa K., Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Wilming L.G., Aidinis V., Allen J.E., Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L., Banal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M., Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R., Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G., Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G., Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M., Georgii-Heminger P., Giegeras T.R., Gojobori T., Green R.E., Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N., Hill D., Humnietek L., Iacono M., Ikeo K., Iwama A., Ishikawa T., Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H., Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K., Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J., Lin S., McWilliam S., Madan Babu M., Madera M., Marchionni L., Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K., Mottagui-Fabriz S., Mulder N., Nakano N., Nakachi H., Ng P., Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O., Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G., Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B., Ringwald M., Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y., Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B., Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K., Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A., Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K., Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C., Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J., Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y., Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T., Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N., Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N., Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S., Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J., Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563 (2005).
RN [5]
- RC TISSUE=Mammary gland;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566 (2005).
RN [6]
- RC TISSUE=Mammary gland;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Oatso N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedziarski R.M., King B.L., Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
- RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [7]
- RC TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [8]
- RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [9]
- RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [10]
- RC TISSUE=Mammary gland;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K., Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S., Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N., Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D., Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A., Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.


```
CC CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC CC Distributed under the Creative Commons Attribution-NoDerivs License
CC CC -----
DR DR EMBL; BC042422; AAH42422.1; -; mRNA.
DR DR EMBL; AK145202; BAE26293.1; -; mRNA.
DR DR HSSP; P07996; 1LSL.

Query Match 100.0%; Score 40; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 210 GVLQNVRF 217

RESULT 12
Q4RLR5 TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DE Chromosome 10 SCAF15019, whole genome shotgun sequence. (Fragment).
DE ORFNames=GSTENG0032374003.
GN Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorphs; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC CC Distributed under the Creative Commons Attribution-NoDerivs License
CC CC -----
DR DR EMBL; CAAB0105019; CAG10667.1; -; Genomic_DNA.
DR DR SMR; Q4RLR5; 834-887, 837-1171.
DR DR GO; GO:0005576; C:extracellular region; IEA.
DR DR GO; GO:0005509; F:calcium ion binding; IEA.
DR DR GO; GO:0008201; F:heparin binding; IEA.
DR DR GO; GO:0005515; F:protein binding; IEA.
DR DR GO; GO:0005198; F:structural molecule activity; IEA.
DR DR GO; GO:0007155; P:cell adhesion; IEA.
DR DR InterPro; IPR006210; EGF.
DR DR InterPro; IPR000742; EGF_3.
DR DR InterPro; IPR01881; EGF_Ca_bd.
DR DR InterPro; IPR013032; EGF_like_reg.
```

```
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Cell adhesion.
FT NON_TER 1171 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 209 GVLQNVRF 216

RESULT 13
TSP1_XENLA STANDARD; PRT; 1173 AA.
ID TSP1_XENLA
DC P35448;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 54.
DE Thrombospondin-1 precursor.
DE Name=thb1; Synonyms=tspl;
GN Xenopus laevis (African clawed frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Ury L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RT "Cloning, characterization and expression of thrombospondin-1 in
RT Xenopus laevis embryos.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-v/beta-1, alpha-
CC v/beta-3 and alpha-11b/beta-3 (by similarity).
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWF domain.
CC CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC CC Distributed under the Creative Commons Attribution-NoDerivs License
CC CC -----
DR DR EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
DR DR HSSP; P07996; 1LSL.
DR DR SMR; P35448; 552-1172.
DR DR InterPro; IPR013320; ConA_like_subgrp.
DR DR InterPro; IPR006210; EGF.
```



```
DR InterPro; IPR006210; EGF
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR013012; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
KW Cell adhesion.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1193;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
DB 189 GVLQNVRF 196

RESULT 15
Q59E99 HUMAN
ID Q59E99 HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBSJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB209912; BAD93149.1; --; mRNA.
DR SMR; Q59E99; 886-939, 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF_3.
DR InterPro; IPR000742; EGF_Ca_bd.
DR InterPro; IPR001881; EGF_3.

DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; TSP_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16E57157B12 CRC64;

Query Match 100.0%; Score 40; DB 2; Length 1225;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
DB 265 GVLQNVRF 272

Search completed: June 5, 2006, 22:42:45
Job time : 82.6983 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 17.9224 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-22

Perfect score: 40

Sequence: 1 XGVQNVRF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pap.*
2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pap.*
3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pap.*
4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pap.*
5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PTUS_COMB.pap.*
6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pap.*
7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	825	2	US-09-949-002-482
2	40	100.0	831	2	US-09-939-853A-97
3	40	100.0	831	2	US-09-939-853A-98
4	40	100.0	1170	1	US-08-313-288B-20
5	40	100.0	1170	2	US-09-657-472-2
6	40	100.0	1170	2	US-09-949-002-350
7	33	82.5	731	2	US-09-270-767-42057
8	32	80.0	484	2	US-09-248-796A-16180
9	31	77.5	57	2	US-09-621-976-5916
10	31	77.5	156	2	US-09-538-092-752
11	31	77.5	238	7	5405943-2
12	31	77.5	400	2	US-09-489-039A-11916
13	31	77.5	406	2	US-09-543-681A-6072
14	31	77.5	436	7	5405943-4
15	31	77.5	911	1	US-08-928-692-59
16	31	77.5	911	2	US-09-339-972-59
17	31	77.5	916	1	US-08-928-692-58
18	31	77.5	916	2	US-09-339-972-58
19	30	75.0	75	2	US-09-248-796A-25342
20	30	75.0	604	2	US-10-094-749-3042
21	29	72.5	229	2	US-09-631-616-15
22	29	72.5	415	2	US-09-198-956-6
23	29	72.5	415	2	US-09-670-141-6
24	29	72.5	416	2	US-09-270-767-43314
25	29	72.5	465	2	US-09-769-787-164
26	29	72.5	561	2	US-09-252-991A-20870

27	29	72.5	609	2	US-09-396-149-9	Sequence 9, Appli
28	29	72.5	715	1	US-08-484-933B-10	Sequence 10, Appl
29	29	72.5	715	1	US-08-484-158B-10	Sequence 10, Appl
30	29	72.5	715	1	US-08-484-596A-10	Sequence 10, Appl
31	29	72.5	715	1	US-08-480-150A-10	Sequence 10, Appl
32	29	72.5	715	2	US-08-458-731-10	Sequence 10, Appl
33	29	72.5	715	2	US-08-149-223A-10	Sequence 10, Appl
34	29	72.5	777	2	US-09-733-643B-15	Sequence 15, Appl
35	29	72.5	1350	1	US-08-319-866-9	Sequence 9, Appli
36	29	72.5	1350	2	US-08-809-917-9	Sequence 9, Appli
37	29	72.5	1350	2	US-09-419-371-9	Sequence 9, Appli
38	28	70.0	110	2	US-09-513-999C-5053	Sequence 5053, Ap
39	28	70.0	120	2	US-09-270-767-46705	Sequence 46705, A
40	28	70.0	166	2	US-08-679-006-31	Sequence 31, Appl
41	28	70.0	166	2	US-09-398-405-31	Sequence 31, Appl
42	28	70.0	172	2	US-09-107-532A-4310	Sequence 4310, Ap
43	28	70.0	172	2	US-09-107-532A-6898	Sequence 6898, Ap
44	28	70.0	172	2	US-09-270-767-33134	Sequence 33134, A
45	28	70.0	172	2	US-09-270-767-48351	Sequence 48351, A

ALIGNMENTS

RESULT 1

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00790
; CURRENT APPLICATION NUMBER: US/09/949.002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825

; TYPE: PRT
; ORGANISM: Human

US-09-949-002-482

Query Match 100.0%; Score 40; DB 2; Length 825;

Best Local Similarity 100.0%; Pred. No. 2.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9

Db 310 GVLQNVRF 317

RESULT 2

US-09-939-853A-97

; Sequence 97, Application US/09939853A

; Patent No. 6989232

; GENERAL INFORMATION:

; APPLICANT: Burgess et al.

; TITLE OF INVENTION: NO. 6989232el Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-099

; CURRENT APPLICATION NUMBER: US/09/939.853A

; CURRENT FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: 60/228,191

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: 60/267,300

; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/269,961

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/277,337

; PRIOR FILING DATE: 2001-03-20

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match      100.0%; Score 40; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232a1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match      100.0%; Score 40; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match      100.0%; Score 40; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match      100.0%; Score 40; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
; US-09-949-002-350

Query Match      100.0%; Score 40; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217
|||||

RESULT 7
US-09-270-767-42057
; Sequence 42057, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42057
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-42057

Query Match      82.5%; Score 33; DB 2; Length 731;
Best Local Similarity 87.5%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      239 GVLQNVDF 246
|||||

RESULT 8
US-09-248-796A-16180
; Sequence 16180, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16180
; LENGTH: 484
```

```
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-16180

Query Match      80.0%; Score 32; DB 2; Length 484;
Best Local Similarity 75.0%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      271 GLLQSVRF 278
|||||

RESULT 9
US-09-621-976-5916
; Sequence 5916, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5916
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -15..-1
; NAME/KEY: UNSURE
; LOCATION: 5
; OTHER INFORMATION: Xaa = Cys,Asp,Gly,Tyr
; US-09-621-976-5916

Query Match      77.5%; Score 31; DB 2; Length 57;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      12 GIQNVSF 19
|||||

RESULT 10
US-09-538-092-752
; Sequence 752, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 752
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YOR226C
```

US-09-538-092-752

Query Match 77.5%; Score 31; DB 2; Length 156;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
|::|||:
Db 76 GIIENVRF 83

RESULT 11

5405943-2
; Patent No. 5405943
; APPLICANT: COMINGS, DAVID E.
; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED

; BEHAVIORS

; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA: US/07/562,596
; APPLICATION NUMBER: US/07/562,596
; FILING DATE: 03-AUG-1990

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 125,577
; FILING DATE: 25-NOV-1987
; APPLICATION NUMBER: 271,653
; FILING DATE: 16-NOV-1988
; APPLICATION NUMBER: 410,831
; FILING DATE: 22-SEP-1989

; SEQ ID NO:2:

; LENGTH: 238

5405943-2

Query Match 77.5%; Score 31; DB 7; Length 238;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 8
|::|||:
Db 166 GVLQNVRF 172

RESULT 12

US-09-489-039A-11916
; Sequence 11916, Application US/09489039A
; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 11916

; LENGTH: 400

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-11916

Query Match 77.5%; Score 31; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLNQVRF 9
|::|||:
Db 121 VLENVRF 127

RESULT 13

US-09-543-681A-6072
; Sequence 6072, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 6072

; LENGTH: 406

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-6072

Query Match 77.5%; Score 31; DB 2; Length 406;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLNQVRF 9
|::|||:
Db 127 VLENVRF 133

RESULT 14

5405943-4

; Patent No. 5405943

; APPLICANT: COMINGS, DAVID E.

; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED

; BEHAVIORS

; NUMBER OF SEQUENCES: 5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/562,596

; FILING DATE: 03-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 125,577

; FILING DATE: 25-NOV-1987

; APPLICATION NUMBER: 271,653

; FILING DATE: 16-NOV-1988

; APPLICATION NUMBER: 410,831

; FILING DATE: 22-SEP-1989

; SEQ ID NO:4:

; LENGTH: 436

5405943-4

Query Match 77.5%; Score 31; DB 7; Length 436;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 8
|::|||:
Db 166 GVLQNVRF 172

RESULT 15

US-08-928-692-59

; Sequence 59, Application US/08928692

; Patent No. 5958727

; GENERAL INFORMATION:

; APPLICANT: Brody, Howard

; APPLICANT: Yaver, Deborah S.

; APPLICANT: Lamsa, Michael

; APPLICANT: Hansen, Kim

; TITLE OF INVENTION: Methods for Modifying the Production of

; TITLE OF INVENTION: a Polypeptide

; NUMBER OF SEQUENCES: 80

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEPT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 911 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 59S8727e
; US-08-928-692-59

Query Match 77.5%; Score 31; DB 1; Length 911;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GVLQNR 8
||:|:|
Db 236 GVMQNR 242

Search completed: June 5, 2006, 22:48:56
Job time : 18.9224 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 58.9655 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-22

Perfect score: 40

Sequence: 1 GVLQNVRF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	12	4	US-10-474-213-28
2	40	100.0	240	4	Sequence 28, Appl
3	40	100.0	240	5	Sequence 40, Appl
4	40	100.0	432	5	Sequence 1020, Ap
5	40	100.0	432	5	Sequence 1022, Ap
6	40	100.0	459	6	Sequence 462, App
7	40	100.0	466	3	Sequence 1047, Ap
8	40	100.0	555	6	Sequence 454, App
9	40	100.0	578	6	Sequence 456, App
10	40	100.0	685	6	Sequence 452, App
11	40	100.0	804	6	Sequence 453, App
12	40	100.0	828	6	Sequence 455, App
13	40	100.0	831	3	Sequence 97, Appl
14	40	100.0	831	3	Sequence 98, Appl
15	40	100.0	855	6	Sequence 461, App
16	40	100.0	1000	6	Sequence 457, App
17	40	100.0	1105	6	Sequence 458, App
18	40	100.0	1150	4	Sequence 1, Appli
19	40	100.0	1152	3	Sequence 1, Appli
20	40	100.0	1169	5	Sequence 7, Appli
21	40	100.0	1170	4	Sequence 12, Appl
22	40	100.0	1170	4	Sequence 2, Appli
23	40	100.0	1170	4	Sequence 114, App
24	40	100.0	1170	4	Sequence 2, Appli
25	40	100.0	1170	4	Sequence 1170, Ap
26	40	100.0	1170	4	Sequence 38, Appl
27	40	100.0	1170	4	Sequence 482, App

28	40	100.0	1170	4	US-10-419-462-38	Sequence 38, Appl
29	40	100.0	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	40	100.0	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	40	100.0	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	40	100.0	1170	5	US-10-782-968-38	Sequence 38, Appl
33	40	100.0	1170	5	US-10-849-989-44	Sequence 44, Appl
34	40	100.0	1170	5	US-10-631-467-548	Sequence 548, App
35	40	100.0	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	40	100.0	1170	5	US-10-831-997-2	Sequence 2, Appli
37	40	100.0	1170	5	US-10-995-561-594	Sequence 594, App
38	40	100.0	1170	5	US-10-995-561-595	Sequence 595, App
39	40	100.0	1170	5	US-10-995-561-596	Sequence 596, App
40	40	100.0	1170	6	US-11-037-713-51	Sequence 51, Appl
41	40	100.0	1170	6	US-11-046-644-28	Sequence 28, Appl
42	40	100.0	1170	6	US-11-046-456-28	Sequence 28, Appl
43	36	90.0	1457	3	US-09-772-316-2	Sequence 2, Appli
44	36	90.0	1457	4	US-10-600-058-2	Sequence 2, Appli
45	36	90.0	1463	5	US-10-750-409-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-10-474-213-28

; Sequence 28, Application US/10474213

; Publication No. US20040214248A1

; GENERAL INFORMATION:

; APPLICANT: Roberts, David D

; APPLICANT: Krutzsch, Henry C

; TITLE OF INVENTION: USE OF SEMENOGLIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF

; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 224329

; CURRENT APPLICATION NUMBER: US/10/474,213

; CURRENT FILING DATE: 2003-10-06

; PRIOR APPLICATION NUMBER: PCT/US02/10535

; PRIOR FILING DATE: 2002-04-03

; PRIOR APPLICATION NUMBER: 60/281,994

; PRIOR FILING DATE: 2001-04-06

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 28

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic peptide

US-10-474-213-28

Query Match 100.0%; Score 40; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.088;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9

|||||||

Db 3 GVLQNVRF 10

RESULT 2

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Williams, Kevin J.

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 40

; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match 100.0%; Score 40; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 192 GVLQNVRF 199
|||||

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; PRIOR FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match 100.0%; Score 40; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 192 GVLQNVRF 199
|||||

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 100.0%; Score 40; DB 5; Length 432;

Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 210 GVLQNVRF 217
|||||

RESULT 5
US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 100.0%; Score 40; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 210 GVLQNVRF 217
|||||

RESULT 6
US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match 100.0%; Score 40; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 210 GVLQNVRF 217
|||||

RESULT 7
US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 100.0%; Score 40; DB 3; Length 466;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
Db 263 GVLQNVRF 270

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 100.0%; Score 40; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
Db 210 GVLQNVRF 217

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 100.0%; Score 40; DB 6; Length 578;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
Db 210 GVLQNVRF 217

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 100.0%; Score 40; DB 6; Length 685;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
Db 210 GVLQNVRF 217

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 100.0%; Score 40; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9
Db 210 GVLQNVRF 217

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match      100.0%; Score 40; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match      100.0%; Score 40; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-09-939-853A-98

Query Match      100.0%; Score 40; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match      100.0%; Score 40; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 GVLQNVRF 9
Db      210 GVLQNVRF 217

Search completed: June 6, 2006, 00:00:11
Job time : 58.9655 secs
```

GenCore version 5.1.9
Copyright (C) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 2.7931 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-22

Perfect score: 40

Sequence: 1 XGVQNVRF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	75.0	608	7	US-11-203-828-5
2	29	72.5	525	6	US-10-953-349-31207
3	29	72.5	560	6	US-10-953-349-31206
4	28	70.0	175	6	US-10-511-937-3000
5	28	70.0	202	6	US-10-953-349-13367
6	28	70.0	212	6	US-10-953-349-13366
7	28	70.0	233	6	US-10-953-349-13365
8	28	70.0	298	6	US-10-953-349-9347
9	28	70.0	308	6	US-10-953-349-16946
10	28	70.0	309	6	US-10-953-349-33145
11	28	70.0	401	6	US-10-953-349-16945
12	28	70.0	402	6	US-10-953-349-33144
13	28	70.0	435	6	US-10-953-349-16944
14	28	70.0	436	6	US-10-953-349-33143
15	27	67.5	329	6	US-10-953-349-23958
16	27	67.5	386	6	US-10-953-349-23720
17	27	67.5	480	6	US-10-953-349-23719
18	27	67.5	528	6	US-10-953-349-23718
19	27	67.5	570	6	US-10-953-349-32501
20	27	67.5	577	6	US-10-953-349-32500
21	27	67.5	581	6	US-10-953-349-32499
22	27	67.5	701	7	US-11-293-697-2701
23	27	67.5	870	7	US-11-316-132-4
24	26	65.0	139	7	US-11-293-697-3323
25	26	65.0	145	6	US-10-953-349-22188

ALIGNMENTS

RESULT 1

US-11-203-828-5
; Sequence 5, Application US/11203828
; Publication No. US20060110390A1
; GENERAL INFORMATION:
; APPLICANT: LEINWAND, LESLIE
; APPLICANT: SUCHAROV, CARMEN
; TITLE OF INVENTION: INHIBITION OF KU AS A TREATMENT FOR CARDIOVASCULAR
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: MYOG:58US
; CURRENT APPLICATION NUMBER: US/11/203,828
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/604,435
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-203-828-5

26 65.0 227 6 US-10-953-349-39138 Sequence 39138, A
26 65.0 262 6 US-10-953-349-2388 Sequence 2388, Ap
26 65.0 276 6 US-10-953-349-39137 Sequence 39137, A
26 65.0 296 6 US-10-953-349-39136 Sequence 39136, A
26 65.0 302 6 US-10-953-349-2387 Sequence 2387, Ap
26 65.0 382 6 US-10-953-349-30058 Sequence 30058, A
26 65.0 387 7 US-11-293-697-3087 Sequence 3087, Ap
26 65.0 408 6 US-10-953-349-30057 Sequence 30057, A
26 65.0 411 6 US-10-953-349-30056 Sequence 30056, A
26 65.0 437 6 US-10-504-120-32 Sequence 32, Appl
26 65.0 440 6 US-10-953-349-1348 Sequence 1348, Ap
26 65.0 459 6 US-10-953-349-7070 Sequence 7070, Ap
26 65.0 470 6 US-10-953-349-7069 Sequence 7069, Ap
26 65.0 557 7 US-11-316-521-39 Sequence 39, Appl
26 65.0 855 7 US-11-247-437-2 Sequence 2, Appl
26 65.0 1033 7 US-11-121-154-207 Sequence 207, Appl
26 65.0 4051 6 US-10-501-834-7 Sequence 7, Appl
26 65.0 4074 6 US-10-501-834-2 Sequence 2, Appl
25.5 63.7 288 6 US-10-953-349-6694 Sequence 6694, Ap
25.5 63.7 318 6 US-10-953-349-6693 Sequence 6693, Ap

RESULT 2

US-10-953-349-31207
; Sequence 31207, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31207
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-31207

Query Match 75.0%; Score 30; DB 7; Length 608;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
Db 463 IVQNVRF 469
:::|||||

RESULT 2

US-10-953-349-31207
; Sequence 31207, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31207
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-31207

US-10-953-349-31207

Query Match 72.5%; Score 29; DB 6; Length 525;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
||| |||
Db 93 GVLDRVRF 100

RESULT 3

US-10-953-349-31206
; Sequence 31206, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31206
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-31206

Query Match 72.5%; Score 29; DB 6; Length 560;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
||| |||
Db 128 GVLDRVRF 135

RESULT 4

US-10-511-937-3000
; Sequence 3000, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3000
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-3000

Query Match 70.0%; Score 28; DB 6; Length 175;
Best Local Similarity 85.7%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
||| |||
Db 36 VLQNVAF 42

RESULT 5

US-10-953-349-13367
; Sequence 13367, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13367
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: misc feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (165)..(165)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-13367

Query Match 70.0%; Score 28; DB 6; Length 202;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNV 7
||| |||
Db 140 GVLQNI 145

RESULT 6

US-10-953-349-13366
; Sequence 13366, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13366
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: misc feature
; LOCATION: (121)..(121)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (138)..(138)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (138)..(138)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (175)..(175)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-13366

Query Match 70.0%; Score 28; DB 6; Length 212;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNV 7
Db 150 GVLQNI 155
|||||

RESULT 7
US-10-953-349-13365
; Sequence 13365, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13365
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142)..(142)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (159)..(159)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (196)..(196)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-13365

Query Match 70.0%; Score 28; DB 6; Length 233;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNV 7
Db 171 GVLQNI 176
|||||

RESULT 8
US-10-953-349-9347
; Sequence 9347, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9347

Query Match 70.0%; Score 28; DB 6; Length 298;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9
Db 88 GLIQGVKF 95
|:|:|:|

RESULT 9
US-10-953-349-16946
; Sequence 16946, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16946
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16946

Query Match 70.0%; Score 28; DB 6; Length 308;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
Db 24 LLENVRF 30
|:|:|:|

RESULT 10
US-10-953-349-33145
; Sequence 33145, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33145
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33145

Query Match 70.0%; Score 28; DB 6; Length 309;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
Db 24 LLENVRF 30
|:|:|:|

RESULT 11
US-10-953-349-16945
; Sequence 16945, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16945
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-16945

; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16945
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16945

Query Match 70.0%; Score 28; DB 6; Length 401;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
:|:||||
Db 117 LLENVRF 123

RESULT 12
US-10-953-349-33144
; Sequence 33144, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33144
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33144

Query Match 70.0%; Score 28; DB 6; Length 402;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
:|:||||
Db 117 LLENVRF 123

RESULT 13
US-10-953-349-16944
; Sequence 16944, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16944
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16944

Query Match 70.0%; Score 28; DB 6; Length 435;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
:|:||||

Db 151 LLENVRF 157

RESULT 14
US-10-953-349-33143
; Sequence 33143, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33143
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33143

Query Match 70.0%; Score 28; DB 6; Length 436;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLQNVRF 9
:|:||||
Db 151 LLENVRF 157

RESULT 15
US-10-953-349-23958
; Sequence 23958, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23958
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-23958

Query Match 67.5%; Score 27; DB 6; Length 329;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 8
:|:||||
Db 193 GTLQNVRF 199

Search completed: June 6, 2006, 00:12:57
Job time : 3.8931 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-24

Perfect score: 62

Sequence: 1 FQGVLLNNRVFVP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq 8:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*
 - 9: Geneseqp2005s:*
 - 10: Geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	12	4	AAB35371
2	56	90.3	12	4	AAB35352
3	56	90.3	12	4	AAB35378
4	56	90.3	12	6	ABG72834
5	56	90.3	240	8	ADL70641
6	56	90.3	432	8	ADQ39359
7	56	90.3	432	8	ADQ39357
8	56	90.3	459	4	AAU02916
9	56	90.3	466	3	ABA43602
10	56	90.3	546	4	AAU02915
11	56	90.3	548	7	ADN02474
12	56	90.3	555	4	AAU02914
13	56	90.3	731	4	AAU02913
14	56	90.3	1152	3	ABO00042
15	56	90.3	1152	5	ABU74771
16	56	90.3	1152	5	ABR82285
17	56	90.3	1170	4	ABR74450
18	56	90.3	1170	4	ABR90800
19	56	90.3	1170	5	AAE25030
20	56	90.3	1170	5	AAU75315
21	56	90.3	1170	6	ABP96780
22	56	90.3	1170	6	ABU03474
23	56	90.3	1170	6	ABG74673

24	56	90.3	1170	6	AAB36228
25	56	90.3	1170	7	ABR62059
26	56	90.3	1170	7	ADN39852
27	56	90.3	1170	8	ADJ76124
28	56	90.3	1170	8	ADJ75296
29	56	90.3	1170	8	ADL70639
30	56	90.3	1170	8	ADL35874
31	56	90.3	1170	8	ADQ26070
32	56	90.3	1170	8	ADP54179
33	56	90.3	1170	8	ADQ39358
34	56	90.3	1170	8	ADQ39356
35	56	90.3	1170	8	ADQ39355
36	56	90.3	1170	9	AD221688
37	56	90.3	1170	9	ABR87781
38	56	90.3	1170	9	ABR46751
39	54	87.1	12	4	AAB35362
40	53	85.5	12	4	AAB35373
41	53	85.5	12	4	AAB35381
42	52	83.9	12	4	AAB35364
43	52	83.9	12	4	AAB35374
44	51	82.3	12	4	AAB35368
45	51	82.3	12	4	AAB35376

ALIGNMENTS

RESULT 1

AAAB35371	
ID	AAB35371 standard; peptide; 12 AA.
XX	
AC	AAB35371;
XX	
DT	08-MAY-2001 (first entry)
XX	
DE	Alpha3beta1 integrin binding peptide #36.
XX	
KW	Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW	diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW	macular degeneration; psoriasis; cell adhesion; cell motility.
XX	
OS	Synthetic.
XX	
PN	WO200105812-A2.
XX	
PD	25-JAN-2001.
XX	
PF	12-JUL-2000; 2000WO-US018986.
XX	
PR	15-JUL-1999; 99US-0144549P.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Roberts DD, Krutzsch HC;
XX	
DR	WPI; 2001-182656/18.
XX	
PT	New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT	useful for inhibiting cell adhesion to extracellular matrix, cell
PT	motility and proliferation and for treating rheumatoid arthritis and
PT	cancer.
XX	
PS	Claim 4; Page 34; 84pp; English.
XX	
CC	The present invention provides a number of peptides which bind to
CC	alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC	and motility, and in the treatment of cancer, diabetic retinopathy,
CC	rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC	and restenosis. The present sequence is an example of one of the peptides
XX	
SQ	Sequence 12 AA;

```
Query Match      100.0%; Score 62; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFVF 12
Db      1 FQGVLLNNRVFVF 12

RESULT 2
AAB35352
ID      AAB35352 standard; peptide; 12 AA.
XX
AC      AAB35352;
XX
DT      08-MAY-2001 (first entry)
XX
DE      Alpha3beta1 integrin binding peptide #17.
XX
KW      Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW      diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW      macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS      Synthetic.
XX
PN      WO200105812-A2.
XX
PD      25-JAN-2001.
XX
PF      12-JUL-2000; 2000WO-US018986.
XX
PR      15-JUL-1999; 99US-0144549P.
XX
PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI      Roberts DD, Kruttsch HC;
XX
DR      WPI; 2001-182656/18.
XX
PT      New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT      useful for inhibiting cell adhesion to extracellular matrix, cell
PT      motility and proliferation and for treating rheumatoid arthritis and
PT      cancer.
XX
PS      Example 2; Page 34; 84pp; English.
XX
CC      The present invention provides a number of peptides which bind to
CC      alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC      and motility, and in the treatment of cancer, diabetic retinopathy,
CC      rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC      and restenosis. The present sequence is an example of one of the peptides
CC      of the invention
XX
SQ      Sequence 12 AA;

Query Match      90.3%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.001;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFVF 12
Db      1 FQGVLLNNRVFVF 12

RESULT 3
AAB35378
ID      AAB35378 standard; peptide; 12 AA.
XX
AC      AAB35378;
XX
DT      08-MAY-2001 (first entry)
XX
DE      Alpha3beta1 integrin binding peptide #43.
XX
```

```
XX      Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW      diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW      macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS      Synthetic.
XX
PN      WO200105812-A2.
XX
PD      25-JAN-2001.
XX
PF      12-JUL-2000; 2000WO-US018986.
XX
PR      15-JUL-1999; 99US-0144549P.
XX
PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI      Roberts DD, Kruttsch HC;
XX
DR      WPI; 2001-182656/18.
XX
PT      New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT      useful for inhibiting cell adhesion to extracellular matrix, cell
PT      motility and proliferation and for treating rheumatoid arthritis and
PT      cancer.
XX
PS      Example 2; Page 34; 84pp; English.
XX
CC      The present invention provides a number of peptides which bind to
CC      alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC      and motility, and in the treatment of cancer, diabetic retinopathy,
CC      rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC      and restenosis. The present sequence is an example of one of the peptides
CC      of the invention
XX
SQ      Sequence 12 AA;

Query Match      90.3%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.001;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFVF 12
Db      1 FQGVLLNNRVFVF 12

RESULT 4
ABG72834
ID      ABG72834 standard; peptide; 12 AA.
XX
AC      ABG72834;
XX
DT      24-FEB-2003 (first entry)
XX
DE      Thrombospondin-1 sequence containing synthetic peptide.
XX
KW      Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW      epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW      colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
OS      Synthetic.
XX
PN      WO200281630-A2.
XX
PD      17-OCT-2002.
XX
PF      03-APR-2002; 2002WO-US010535.
XX
PR      06-APR-2001; 2001US-0281994P.
XX
PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI      Roberts DD, Kruttsch HC;
```

XX WPI; 2003-103329/09.

XX A new diagnosis for cancer other than prostate cancer in a mammal useful

PT to detect cancer including lung cancer, particularly small cell lung

PT cancer and melanoma comprises detecting semenogelin in a sample.

XX

PS Example 1; Page 14; 32pp; English.

XX

CC The invention relates to diagnosing cancer other than prostate cancer in

CC a male mammal, comprising assaying a test sample for increased level of

CC semenogelin, or cancer in a female by assaying for the presence of

CC semenogelin. Administering a semenogelin protein or polypeptide fragment

CC or a semenogelin-specific antibody or active fragment, or a recombinant

CC vector expressing the protein or antibody, is useful for inducing an

CC immune response to a cancer in a mammal, where the cancer is not prostate

CC cancer and semenogelin is a marker. The invention is used to diagnose

CC cancer, particularly of epithelial origin such as lung cancer, papillary

CC renal cell carcinoma, colon cancer, especially small-cell lung cancer

CC (SCLC), or a melanoma. The present sequence represents the amino acid

CC sequence of the thrombospondin-1 sequence containing synthetic peptide

CC which binds to alpha-3-beta-1 integrin

XX

XX Sequence 12 AA;

Query Match 90.3%; Score 56; DB 6; Length 12;

Best Local Similarity 91.7%; Pred. No. 0.001;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLNVRVVF 12

DB 1 FQGVLNVRVVF 12

RESULT 5

ADL70641

ID ADL70641 standard; protein; 240 AA.

XX

AC ADL70641;

XX

XX 20-MAY-2004 (first entry)

XX

DE Human thrombospondin-1 N-terminal domain.

XX

XX Human; thrombospondin-1; epitope; cancer; diagnosis.

XX

OS Homo sapiens.

PH Key Location/Qualifiers

FT Region 23..32

FT /note= "Heparin binding region"

FT Region 77..82

FT /note= "Heparin binding region"

FT Region 151..164

FT /note= "Fibrinogen binding region"

XX

PN W02004018995-A2.

XX

XX 04-MAR-2004.

XX

XX 20-AUG-2003; 2003WO-US026023.

XX

XX 23-AUG-2002; 2002US-0405494P.

XX

XX 21-APR-2003; 2003US-00419462.

XX

XX (WILL/) WILLIAMS K J.

XX

XX Williams KJ;

PI

XX WPI; 2004-226901/21.

XX

XX New purified thrombospondin fragment extracted from a body fluid, useful

PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma

PT or leukemia or as calibrators, indicators, immunogens and analytes.

XX

PS Disclosure; SEQ ID NO 40; 76pp; English.

XX

CC The present sequence is that of the N-terminal domain of human

CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments

CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in

CC clinical assays for cancer and for generation of antibodies and other

CC binding agents. A method that distinguishes TSP from a TSP fragment or

CC portion involves: (1) using an epitope shared by TSP and the TSP fragment

CC or portion as a target for a binding molecule, e.g. an antibody, to

CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an

CC epitope present in TSP but not in the fragment or portion to obtain a

CC quantitation of TSP only; and (3) using the difference between (1) and

CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable

CC epitopes are provided ADL70602-ADL70638. Detection or quantification of

CC the TSP fragment or portion is performed in order to detect the presence,

CC or monitor the course, of a disease or condition selected from cancer,

CC renal failure, renal disease, atopic dermatitis, vasculitis, acute

CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial

CC infarction, liver disease, splenectomy, dermatomyositis, polyarteritis

CC nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki

CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,

CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,

CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition

CC associated with clotting, a condition associated with platelet

CC activation, a condition associated with intravascular platelet

CC activation, a condition associated with consumption of platelets, heparin

CC -induced thrombocytopenia, disseminated intravascular coagulation,

CC intravascular coagulation, extravascular coagulation, a condition

CC associated with endothelial activation, a condition associated with

CC production and/or release of thrombospondin and/or a thrombospondin

CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic

CC reaction, an aspartame reaction, atopic dermatitis, eczema,

CC hypersensitivity, scleroderma, conditions associated with plugging of

CC vessels, a condition associated with a cryofibrinogen, a condition

CC associated with a cryoglobulin, and a condition associated with an anti-

CC cardiolin antibody. The cancer is selected from adenoma,

CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid

CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,

CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of

CC the respiratory system, circulatory system, musculoskeletal system,

CC muscle, bone, a joint, tendon or ligament, digestive system, liver or

CC biliary system, pancreas, head, neck, endocrine system, reproductive

CC system (male or female), genitourinary system, kidney, urinary tract,

CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary

CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal

CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a

CC well-differentiated cancer or a moderately differentiated cancer.

XX

XX Sequence 240 AA;

Query Match 90.3%; Score 56; DB 8; Length 240;

Best Local Similarity 91.7%; Pred. No. 0.029;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLNVRVVF 12

DB 190 FQGVLNVRVVF 201

RESULT 6

ADQ39359

ID ADQ39359 standard; protein; 432 AA.

XX

XX ADQ39359;

XX

XX 18-NOV-2004 (first entry)

XX

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

DE

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX

XX cardiant; gene therapy; human.

XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR N-PSDB; ADQ38531.
XX PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1022; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;
Query Match 90.3%; Score 56; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLNVRVVF 12
Db 208 FQGVLNVRVVF 219
RESULT 7
ADQ39357
ID ADQ39357 standard; protein; 432 AA.
XX AC ADQ39357;
XX DT 18-NOV-2004 (first entry)

XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1020.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human.
XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR N-PSDB; ADQ38529.
XX PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1020; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC polynucleotide, a buffer and an enzyme; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;
Query Match 90.3%; Score 56; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLNVRVVF 12
Db 208 FQGVLNVRVVF 219
RESULT 8
AAU02916

ID AAU02916 standard; protein; 459 AA.
 AC AAU02916;
 XX
 XX
 DT 12-SEP-2001 (first entry)
 DE
 DE Angiotensin converting enzyme (ACEV) splice variant protein #16.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW noncardioidic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO200136632-A2.
 XX
 XX 25-MAY-2001.
 XX
 XX 17-NOV-2000; 2000WO-IL000766.
 XX
 XX 17-NOV-1999; 99IL-00132978.
 PR 10-DEC-1999; 99IL-00133455.
 XX
 XX (COMP-) COMPUGEN LTD.
 PA
 XX
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI
 XX WPI; 2001-336004/35.
 DR N-PSDB; AAS06016.
 DR
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 XX Claim 4; Fig 16; 519pp; English.
 PS
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, noncardioidic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 459 AA;
 Query Match 90.3%; Score 56; DB 4; Length 459;
 Best Local Similarity 91.7%; Pred. No. 0.059;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVLNVRVVF 12
 |||||
 DB 208 FQGVLNVRVVF 219
 RESULT 9
 AAB43602

ID AAB43602 standard; protein; 466 AA.
 AC AAB43602;
 XX
 XX
 DT 08-FEB-2001 (first entry)
 DE
 DE Human cancer associated protein sequence SEQ ID NO:1047.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatologic; neuroprotective; thrombolytic; coagulant; inotropic;
 KW vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 XX WO200055350-A1.
 PN
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US005882.
 XX
 XX 12-MAR-1999; 99US-0124270P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587533/55.
 DR N-PSDB; AAC77811.
 DR
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 PT
 XX Claim 11; Page 1636-1638; 2352pp; English.
 PS
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cycostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
 CC vasotropic; antiporiatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 466 AA;
 Query Match 90.3%; Score 56; DB 3; Length 466;
 Best Local Similarity 91.7%; Pred. No. 0.06;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVLNVRVVF 12
 |||||
 DB 261 FQGVLNVRVVF 272

RESULT 10
AAU02915
ID AAU02915 standard; protein; 546 AA.
XX AC
XX AAU02915;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX
OS Homo sapiens.
XX
PN WO200136632-A2.
XX
PD 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-IL000766.
XX
PF 17-NOV-1999; 99IL-00132978.
XX
PR 10-DEC-1999; 99IL-00133455.
XX
XX (COMP-) COMPUGEN LTD.
XX
PA Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX
PI WPI; 2001-336004/35.
XX
DR N-PSDB; AAS06015.
XX
XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 15; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding to the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 546 AA;
XX
Query Match 90.3%; Score 56; DB 4; Length 546;
Best Local Similarity 91.7%; Pred. No. 0.072;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLLNNVRVVF 12
| | | | | | | | | |
Db 208 FQGVLLNNVRVVF 219

RESULT 11
ADN02474
ID ADN02474 standard; protein; 548 AA.
XX AC
XX ADN02474;
XX
DT 17-JUN-2004 (first entry)
XX
DE TSF polypeptide.
XX
KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX
OS Homo sapiens.
XX
PN CN1401387-A.
XX
PD 12-MAR-2003.
XX
PF 21-AUG-2002; 2002CN-00129408.
XX
PR 21-AUG-2002; 2002CN-00129408.
XX
XX (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
XX
XX Han Z, Liu P;
XX
DR WPI; 2003-469302/45.
XX
DR N-PSDB; ADN02475.
XX
XX Tumor suppressing polypeptide TSF and gene therapy vector composition.
XX
PS Claim 2; SEQ ID NO 1; 13pp; Chinese.
XX
XX The present invention relates to a novel recombinant adenovirus vector
CC mediated anti-neoplastic composition is prepared through cloning the cDNA
CC sequence from the human peripheral blood cell by specific primer and
CC reverse transcription-polymerase chain reaction (RT-PCR) method for
CC coding TSF polypeptide, construction in human embryonic kidney cell 293
CC by AdEasy system, and packaging and expressing the recombinant adenovirus
CC vector of TSF. It can suppress the growth and transfer of cancer. The
CC present sequence represents the TSF polypeptide.
XX
SQ Sequence 548 AA;
XX
Query Match 90.3%; Score 56; DB 7; Length 548;
Best Local Similarity 91.7%; Pred. No. 0.072;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLLNNVRVVF 12
| | | | | | | | | |
Db 208 FQGVLLNNVRVVF 219

RESULT 12
AAU02914
ID AAU02914 standard; protein; 555 AA.
XX AC
XX AAU02914;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW

KW vascular disorder; asbestosis.

OS Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

PR 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

DR N-PSDB; AAS06014.

XX Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

PS Claim 4; Fig 14; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis

XX Sequence 555 AA;

Query Match 90.3%; Score 56; DB 4; Length 555;

Best Local Similarity 91.7%; Pred. No. 0.073;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12

Db 208 FQGVLNVRVVF 219

RESULT 13

AAU02913

ID AAU02913 standard; protein; 731 AA.

XX AAU02913;

DT 12-SEP-2001 (first entry)

XX Angiotensin converting enzyme (ACEV) splice variant protein #13.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6; granulocyte colony stimulating factor receptor; glucagon; hypertrophy; platelet-derived endothelial cell growth factor; cardiovascular disease; cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer; myocardial infarction; coronary arterial thrombosis; renal disease; diabetic nephropathy; muscular disease; immune disorder; sarcoidosis; multiple sclerosis; immune complex nephritis; deep vein thrombosis; nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

KW vascular disorder; asbestosis.

OS Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

PR 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

DR N-PSDB; AAS06013.

XX Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

PS Claim 4; Fig 13; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis

XX Sequence 731 AA;

Query Match 90.3%; Score 56; DB 4; Length 731;

Best Local Similarity 91.7%; Pred. No. 0.099;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12

Db 208 FQGVLNVRVVF 219

RESULT 14

AAB00042

ID AAB00042 standard; protein; 1152 AA.

XX AAB00042;

DT 08-NOV-2000 (first entry)

XX Human thrombospondon-1 (TSP-1).

XX TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin; angiogenesis; tumour; treatment; cancer; arthritis; psoriasis; diabetic retinopathy; corneal graft rejection; glaucoma.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Region 361..416

/label= type 1 repeat region

```

FT Region 417. .473
FT /label= Type 1 repeat region
FT Region 474. .530
FT /label= Type 1 repeat region
XX
XX WO200044908-A2.
XX
XX 03-AUG-2000.
XX
XX 01-FEB-2000; 2000WO-US002482.
XX
XX 01-FEB-1999; 99US-0118053P.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX Lawler JW;
XX
XX WPI; 2000-514823/46.
XX
XX Nucleic acids encoding chimeric proteins such as cartilage oligomeric
XX matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
XX inhibiting angiogenesis and treating diseases such as cancer.
XX
XX Disclosure; Fig 1; 40pp; English.
XX
XX New nucleic acids are described which encode a protein comprising the
XX second and third type 1 repeats of human TSP (thrombospondin)-1, but not
XX the TGF (transforming growth factor)-beta activation region of human TSP-
XX 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
XX third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
XX assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
XX reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
XX tumours in mice models. Thus the nucleic acids and proteins may be useful
XX for treating angiogenesis related diseases such as cancer (by reducing
XX the rate of growth and size of tumours), arthritis, psoriasis, diabetic
XX retinopathy, corneal graft rejection, and glaucoma. They may also be used
XX for treating human immunodeficiency virus (HIV) infection. Anti-
XX angiogenic therapy has little toxicity, does not require the therapeutic
XX agent to enter tumour cells or cross the blood-brain barrier, controls
XX tumour growth independently of growth of tumour cell heterogeneity, and
XX does not induce drug resistance
XX
XX Sequence 1152 AA;
XX
XX Query Match 90.3%; Score 56; DB 3; Length 1152;
XX Best Local Similarity 91.7%; Pred. No. 0.16;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLNVRVVF 12
XX ||||| |||||
XX Db 190 FQGVLNVRVVF 201
XX
XX RESULT 15
XX AAU74771
XX ID AAU74771 standard; protein; 1152 AA.
XX
XX AC AAU74771;
XX
XX 09-APR-2002 (first entry)
XX
XX Human thrombospondin-1 (TSP-1).
XX
XX Thrombospondin-1; TSP-1; cytostatic; angiogenesis; vasotropic; vulnery;
XX neovascularisation; cell proliferation inhibitor; cancer; solid tumour;
XX haemangioma; acoustic neuromas; neurofibroma; trachoma;
XX pyogenic granulomas; rheumatoid arthritis; ocular angiogenic disease;
XX retinopathy; psoriasis; macular degeneration; corneal graft rejection;
XX neovascular glaucoma; retrolental fibroplasia; rebesosis; angiofibroma;
XX Osier-Webber syndrome; myocardial angiogenesis; haemophilic joints;
XX plaque neovascularisation; telangiectasia; wound granulation; apoptosis.
XX
XX Homo sapiens.

```

```

XX Key Location/Qualifiers
XX Region 263. .360
XX Domain /label= Procollagen_homology_region
XX 303. .309
XX /label= Procollagen_homology_domain
XX /note= "Required in inhibition of angiogenesis"
XX 361. .530
XX /label= Type 1 repeat domain
XX /note= "This region contains 3 type 1 repeats, from
XX residues 361-416, residues 417-473 and residues 474-530"
XX 364. .370
XX /label= Heparin_binding_domain
XX 413. .415
XX /label= RFX motif
XX /note= "Necessary and sufficient for activation of
XX transforming growth factor beta (TGF beta)"
XX 418. .423
XX /label= TGF-beta and fibronectin binding_domain
XX /note= "Transforming growth factor"
XX 420. .426
XX /label= Heparin_binding_domain
XX 429. .434
XX /label= Cell_binding_domain
XX 477. .483
XX /label= Heparin_binding_domain
XX 481. .499
XX /label= Anti-angiogenesis_domain
XX 486. .491
XX /label= Cell_binding_domain
XX 531. .673
XX /label= Type 2 repeat domain
XX /note= "This region contains 3 type 2 repeats, from
XX residues 531-571, residues 572-629 and residues 630-673"
XX 570. .601
XX /label= Calcium_binding_domain
XX 698. .925
XX /label= Type 3 repeat domain
XX /note= "This region contains 7 type 3 repeats, from
XX residues 698-733, residues 734-756, residues 757-792,
XX residues 793-815, residues 816-853, residues 854-889 and
XX residues 890-925"
XX
XX WO200191781-A2.
XX
XX 06-DEC-2001.
XX
XX 25-MAY-2001; 2001WO-US017250.
XX
XX 26-MAY-2000; 2000US-0207994P.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX Lawler JW;
XX
XX WPI; 2002-106273/14.
XX
XX Composition useful for treatment of cancer comprises cDNA encoding amino
XX acids of human thrombospondin-1 or its conservative variant and a
XX carrier.
XX
XX Disclosure; Fig 7; 54pp; English.
XX
XX The invention describes a composition comprising cDNA encoding fragments
XX of human thrombospondin-1 (TSP -1), a type 1 repeat polypeptide and
XX potent inhibitor of tumour growth and angiogenesis. The composition is
XX useful for killing cancerous cells (preferably tumour); for reducing
XX volume or inhibiting growth of a tumour (inhibiting neovascularisation in
XX the tumour); for decreasing proliferation of tumour cells; in the
XX treatment of diseases and conditions associated with angiogenic activity
XX or misregulated growth and angiogenesis-mediated diseases such as cancer,
XX solid tumour, tumour metastasis, benign tumour, (e.g. haemangioma,
XX acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas),

```

CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases (e.g.
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasias,
CC reboosis), Osler-Webber syndrome, myocardial angiogenesis,
CC telangiectasia, plaque neovascularisation, haemophilic joints,
CC angiofibroma or wound granulation. The composition induces apoptosis and
CC inhibits neovascularisation in the tumour cells. This amino acid sequence
CC represents human thrombospondin-1 (TSP-1), on which the recombinant
CC proteins of the invention are based
XX
SQ Sequence 1152 AA;

Query Match 90.3%; Score 56; DB 5; Length 1152;
Best Local Similarity 91.7%; Pred. No. 0.16;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLNVRVF 12
| | | | |
Db 190 FQGVLNVRVF 201

Search completed: June 5, 2006, 22:24:58
Job time : 94.1379 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-24

Perfect score: 62

Sequence: 1 FQGVLNVRVVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	90.3	229	2	thrombospondin 1 -
2	56	90.3	1170	1	thrombospondin 1 p
3	56	90.3	1170	2	thrombospondin 1 p
4	40	64.5	597	2	hypothetical prote
5	39	62.9	311	2	ribose/galactose A
6	39	62.9	876	2	probable receptor
7	38	61.3	465	2	probable membrane
8	38	61.3	1172	1	thrombospondin 2 p
9	38	61.3	1172	2	thrombospondin 2 p
10	37	59.7	882	2	valyl-tRNA synthet
11	37	59.7	883	2	valyl-tRNA synthet
12	37	59.7	943	2	excinuclease ABC c
13	37	59.7	943	2	excinuclease ABC,
14	36	58.1	385	2	hypothetical prote
15	36	58.1	467	2	H4-transporting tw
16	36	58.1	569	2	laccase-like prote
17	36	58.1	755	2	hypothetical prote
18	36	58.1	1178	1	thrombospondin pre
19	36	58.1	1288	2	5-oxoprolinase (AT
20	36	58.1	1839	2	mannuronan C-5-epi
21	35	56.5	247	1	probable 3-oxoacyl
22	35	56.5	247	2	3-oxoacyl-[acyl-ca
23	35	56.5	271	2	hypothetical prote
24	35	56.5	316	2	hypothetical prote
25	35	56.5	440	2	hypothetical prote
26	35	56.5	565	2	hypothetical prote
27	35	56.5	603	2	sensory transducti
28	35	56.5	834	2	probable integrat
29	35	56.5	1054	2	reverse gyrase (to

30 35 56.5 1071 2 T43255
31 35 56.5 1305 2 A0428
32 35 56.5 1582 2 A56248
33 35 56.5 1829 2 AE1864
34 35 56.5 2018 2 T34274
35 35 56.5 3005 1 GNVSTV
36 35 56.5 4096 2 A57099
37 34 54.8 111 2 S2596
38 34 54.8 133 2 B89969
39 34 54.8 186 2 B75421
40 34 54.8 198 2 B64363
41 34 54.8 212 2 D81929
42 34 54.8 262 2 A75009
43 34 54.8 304 2 T34271
44 34 54.8 306 2 B96922
45 34 54.8 350 2 P89883

tricorn proteinase
probable exported
sulfonylurea recep
hypothetical prote
hypothetical prote
genome polypeptide
DNA-activated prot
probable membrane
enterotoxin ventl
probable pilin, ty
imidazoleglycerol-
probable imidazole
probable aryl phos
hypothetical prote
transcription regu
cell surface prote

ALIGNMENTS

RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C:Accession: S57957

R:Lafuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi
A:Reference number: S57955

A:Accession: S57957

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-229 <LAF>

A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:G899228; PIDN

C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

Query Match 90.3%; Score 56; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0041;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12

||||| |||||

Db 190 FQGVLNVRVVF 201

RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)

C>Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A26155; A34274; A30140; A25812; A05172; A42927

R:Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c

A:Reference number: A26155; MUID:87057617; PMID:2430973

A:Accession: A26155

A:Molecule type: mRNA

A:Residues: 1-1170 <LAW>

A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:G37137; PIDN:CF

A>Note: parts of this sequence, including the amino end of the mature protein, were deter

R:Lahterty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA se

A:Reference number: A34274; MUID:89291870; PMID:2544587

A:Accession: A34274

A:Molecule type: DNA

A:Residues: 1-166 <LAH>

A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835

R:Hennessey, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I

J. Cell Biol. 108, 729-736, 1989

A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

C:Genetics:

A:Gene: UUS19

A:Genetic code: SGC3

C:Superfamily: Ureaplasma urealyticum hypothetical protein UUS20

Query Match

64.5%; Score 40; DB 2; Length 597;

Best Local Similarity 63.6%; Pred. No. 14;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy

2 QGVLNVRVF 12

|:|||||:

Db 397 QGILNPLRV 407

RESULT 5

F70184

ribose/galactose ABC transporter, permease protein (rbsC-2) homolog - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: F70184

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: F70184

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-311 <KLE>

A:Cross-references: UNIPROT:O51622; UNIPARC:UPI000005758D; GB:AE001168; GB:AE000783; NID

A:Experimental source: strain B31

C:Superfamily: probable ribose ABC transporter rbsC-2

Query Match

62.9%; Score 39; DB 2; Length 311;

Best Local Similarity 50.0%; Pred. No. 11;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy

1 FQGVLNVRVF 12

|:|||||:

Db 131 FDGILNKTSTF 142

RESULT 6

B96693

probable receptor serine/threonine kinase PRSK T4O24.2 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: B96693

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B96693

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-876 <STO>

A:Cross-references: UNIPROT:Q9C622; UNIPARC:UPI000004840F; GB:AE005173; NID:gill128393; F

C:Genetics:

A:Gene: T4O24.2

A:Map position: 1

Query Match

62.9%; Score 39; DB 2; Length 876;

Best Local Similarity 77.8%; Pred. No. 33;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLNVR 9

|:|||||:

Db 393 FRGVNLNLR 401

RESULT 7

AC0347

probable membrane protein yegB [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AC0347

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;

geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, E.

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0347

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-465 <KUR>

A:Cross-references: UNIPROT:Q8ZCV8; UNIPARC:UPI00000DC696; GB:AL590842; PIDN:CAC921102.1;

C:Genetics:

A:Gene: yegB

C:Superfamily: multidrug-efflux transporter

Query Match

61.3%; Score 38; DB 2; Length 465;

Best Local Similarity 70.0%; Pred. No. 26;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVR 10

|:|||||:

Db 362 FQGMNLALRF 371

RESULT 8

TSHUP2

thrombospondin 2 precursor - human

C:Species: Homo sapiens (man)

C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A47379; A42173

R:LaBell, T.L.; Byers, P.H.

Genomics 17, 225-229, 1993

A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: poter

A:Reference number: A47379; MUID:94010892; PMID:8406456

A:Accession: A47379

A:Molecule type: mRNA

A:Residues: 1-1172 <LAB>

A:Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046680; GB:L12350; NID:g307505; PIDN:F

R:LaBell, T.L.; Milewicz, D.J.; Distèche, C.M.; Byers, P.H.

Genomics 12, 421-429, 1992

A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression of

A:Reference number: A42173; MUID:92217961; PMID:1559694

A:Accession: A42173

A:Molecule type: mRNA

A:Residues: 560-1172 <LA2>

A:Cross-references: UNIPARC:UPI00001742C1; GB:M81339

A:Experimental source: fibroblast

A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIIP:95096)

C:Genetics:

A:Gene: GDB:THBS2; TSP2

A:Cross-references: GDB:128789; OMIM:188061

A:Map position: 6Q27-6q27

C:Complex: homotrimer, disulfide linked

C:Function:

A:Description: participates in cell migration and adhesion, and in platelet aggregation

C:Superfamily: thrombospondin 1; SGF homology; thrombospondin type 1 repeat homology; vor

C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-1172/Product: thrombospondin 2 #status predicted <MAT>

F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>

F:380-431/Domain: thrombospondin type 1 repeat homology <THRI>

C;Genetics:
A;Gene: uvvRA
C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 59.7%; Score 37; DB 2; Length 943;
Best Local Similarity 55.6%; Pred. No. 87;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLNNVR 9
 |.:|.|.|.:
Db 373 FEGVINNK 381

RESULT 13
F95021
excinuclease ABC, chain A [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95021
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95021
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-943 <KUR>
A;Cross-references: UNIPROT:Q9T5X7; UNIPARC:UPI0000137E9B; GB:AEO05672; PIDN:AAK74367.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0186
C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 59.7%; Score 37; DB 2; Length 943;
 Best Local Similarity 55.6%; Pred. No. 87;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVNNVR 9
 | : | : | : | : | :
 Db 373 FEGVINNIK 381

RESULT 14
 S56224
 hypothetical protein YFI030w - Yeast (*Saccharomyces cerevisiae*)
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 05-Oct-2004
 C:Accession: S56224
 R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanu
 submitted to the EMBL Data Library, May 1995
 A:Description: Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cere*
 A:Reference number: S56186
 A:Accession: S56224
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-385 <MUR>
 A:Cross-references: UNIPROT:P43567; UNIPARC:UPI000013AE1E; EMBL:D50617; NID:G836685; PII:
 C:Genetics:
 A:Cross-references: SGD:S0001864
 A:Map position: 6L
 C:Superfamily: serine-pyruvate/aspartate aminotransferase

```
Query Match      58.1%; Score 36; DB 2; Length 385;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 8: Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 1 FQGVLNVRVF 12
|||
Db 40 FQVLRNTRAVF 51

RESULT 15

D84938
H+-transporting two-sector ATPase (EC 3.6.3.14), flagellum-specific [imported] - Buchnera
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C:Accession: D84938
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. AF
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: D84938
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <STO>
A:Cross-references: UNIPARC:UPI0000005E44F; GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: flil; BU076
C:Superfamily: H(+)-transporting ATP synthase; H+-transporting ATP synthase alpha chain
C:Keywords: hydrolase

Query Match	58.1%	Score 36;	DB 2;	Length 467;
Best Local Similarity	60.0%	Pred. No. 62;		
Matches	6;	Conservative	3;	Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLNNVRF 10
|:|:|:|:|
Db 10 FRGILNNLRF 19

Search completed: June 5, 2006, 22:44:57
Job time : 15.9655 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 108.931 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-24

Perfect score: 62

Sequence: 1 FQGVLLNVRVFV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	56	90.3	229	2	Q28194_BOVIN	Q28194 bos taurus
2	56	90.3	496	2	Q7SY84_XENLA	Q7SY84 xenopus lae
3	56	90.3	1170	1	TSPI_BOVIN	Q28178 bos taurus
4	56	90.3	1170	1	TSPI_HUMAN	P07996 homo sapien
5	56	90.3	1170	1	TSPI_MOUSE	P35441 mus musculus
6	56	90.3	1170	2	Q3TR40_MOUSE	Q3TR40 mus musculus
7	56	90.3	1170	2	Q7LSA3_RAT	Q7LSA3 rattus norv
8	56	90.3	1171	2	Q80YQ1_MOUSE	Q80YQ1 mus musculus
9	56	90.3	1171	2	Q8CGB2_MOUSE	Q8CGB2 mus musculus
10	56	90.3	1173	1	TSPI_XENLA	P35448 xenopus lae
11	56	90.3	1225	2	Q59E99_HUMAN	Q59E99 homo sapien
12	51	82.3	1090	2	Q5SPG5_BRARE	Q5SPG5 brachydanio
13	51	82.3	1193	2	Q4S758_TETNG	Q4S758 tetraodon n
14	50	80.6	249	2	Q5U903_PIG	Q5U903 sus scrofa
15	50	80.6	1171	2	Q4RLR5_TETNG	Q4RLR5 tetraodon n
16	46	74.2	1034	2	Q4RQ74_TETNG	Q4RQ74 tetraodon n
17	44	71.0	727	2	Q8PRY3_METBA	Q8PRY3 methanosarc
18	44	71.0	730	2	Q467V1_METBA	Q467V1 methanosarc
19	44	71.0	733	2	Q8TLX6_METAC	Q8TLX6 methanosarc
20	42	67.7	84	2	Q60XF5_CABBR	Q60XF5 caenorhabdi
21	42	67.7	100	2	Q83E24_COKBU	Q83E24 coxiella bu
22	42	67.7	328	2	Q2LSR0_CLOPE	Q2LSR0 clostridium
23	42	67.7	713	2	Q3FLU8_BURK	Q3FLU8 burkholderi
24	42	67.7	713	2	Q44XL2_BURK	Q44XL2 burkholderi
25	42	67.7	713	2	Q4LLM8_BURK	Q4LLM8 burkholderi
26	40	64.5	302	2	Q4YQ55_PLABE	Q4YQ55 plasmodium
27	40	64.5	597	2	Q9PPX0_UREPA	Q9PPX0 ureaplasma
28	40	64.5	659	2	Q8RC10_THETN	Q8RC10 thermoaer
29	40	64.5	689	2	Q8R808_THETN	Q8R808 thermoaer
30	40	64.5	1034	2	Q4DWH6_TRYCR	Q4DWH6 trypanosoma
31	40	64.5	1034	2	Q7Z2B9_TRYCR	Q7Z2B9 trypanosoma

32	39	62.9	212	2	Q5FA21_NEIG1	Q5FA21 neisseria g
33	39	62.9	311	2	O51622_BORBU	O51622 borrelia bu
34	39	62.9	378	2	Q563V1_XENLA	Q563V1 xenopus lae
35	39	62.9	380	2	Q563S6_9PIPI	Q563S6 xenopus ruw
36	39	62.9	380	2	Q563S7_9PIPI	Q563S7 xenopus ruw
37	39	62.9	380	2	Q563S8_9PIPI	Q563S8 xenopus and
38	39	62.9	380	2	Q563S9_9PIPI	Q563S9 xenopus and
39	39	62.9	380	2	Q563T0_9PIPI	Q563T0 xenopus ruw
40	39	62.9	380	2	Q563T1_9PIPI	Q563T1 xenopus lon
41	39	62.9	380	2	Q563T2_9PIPI	Q563T2 xenopus lon
42	39	62.9	380	2	Q563T3_9PIPI	Q563T3 xenopus lon
43	39	62.9	380	2	Q563T4_9PIPI	Q563T4 xenopus ami
44	39	62.9	380	2	Q563T5_9PIPI	Q563T5 xenopus ami
45	39	62.9	380	2	Q563T6_9PIPI	Q563T6 xenopus pyg

ALIGNMENTS

RESULT 1

ID Q28194_BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96331130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCPI9>3.3.CO;2-0;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License

CC -----
DR EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON_TER 1
FT NON_TER 229 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9BECB4E6B669C CRC64;

Query Match 90.3%; Score 56; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. NO. 0.037;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLLNVRVFV 12

|||||

190 FQGVLLNVRVFV 201

RESULT 2

Q7SY84_XENLA
ID Q7SY84_XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.

DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausnitzer R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole.
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC054970; AAH54970.1; -; mRNA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00090; TSP_1; 2.
CC Pfam; PF00093; VWC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 2.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS50092; TSP1; 2.
CC PROSITE; PS01208; VWF_C_1; UNKNOWN_1.
CC PROSITE; PS0184; VWF_C_2; 1.
SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7BF51B CRC64;

Query Match 90.3%; Score 56; DB 2; Length 496;
Best Local Similarity 91.7%; Pred. No. 0.085;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FQGLVNVRFVF 12
||||| |||||

Db 214 FQGLVNVRFVF 225
RESULT 3
TSPL_BOVIN
ID TSPL_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSPL;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RX [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
odontoblasts and preodontoblasts";
RL Biochim. Biophys. Acta 1382:17-22 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
TGF-beta.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
laminin, type V collagen and integrins alpha-V/beta-1, alpha-
V/beta-3 and alpha-11b/beta-3. May play a role in dentinogenesis
and/or maintenance of dentin and dental pulp.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- TISSUE SPECIFICITY: Odontoblasts.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWF domain.
CC -----
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB005287; BAA21115.1; -; mRNA.
CC EMBL; X87618; CAA60950.1; -; mRNA.
CC EMBL; X87619; CAA60951.1; -; mRNA.
CC PIR; S55501; S55501.
CC HSSP; P07996; 1LSL.
CC SWE; Q28178; 549-1169.
CC GlycoSuiteDB; Q28178; -;
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000742; EGF 3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR003367; tsp 3.
CC InterPro; IPR008859; TSP_C.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 13.

DR Pfam; PF00735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VMC; 1; 1.
DR PROSITE; PS0184; VMC; 2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT REGION 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT CARBOHYD 1085 1085
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT DISULFID 519 531
FT DISULFID 551 562
FT DISULFID 556 572
FT DISULFID 575 586
FT DISULFID 592 608
FT DISULFID 599 617
FT DISULFID 620 644
FT DISULFID 650 663
FT DISULFID 657 676
FT DISULFID 678 689
FT DISULFID 705 713
FT DISULFID 718 738
FT DISULFID 754 774
FT DISULFID 777 797
FT DISULFID 813 833
FT DISULFID 836 856
FT DISULFID 874 894
FT DISULFID 910 930
FT DISULFID 946 1167
FT CONFLICT 805 805 S -> G (in Ref. 2).
SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 90.3%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLLNNRVFVF 12
DB 208 FQGVLLNNRVFVF 219
RESULT 4
ID TSP1 HUMAN STANDARD; PRT; 1170 AA.
AC P07996; Q15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DT 07-MAR-2006, entry version 78.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologues with several different
RT proteins";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=83139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Franson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
TRP-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;

RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.,
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RL module."; J. Biol. Chem. 276:6485-6498 (2001).
RN [8]
RN THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RL of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RN Biochemistry 41:14329-14339 (2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hydrazide chemistry, and mass spectrometry."; J.
RL J. Proteome Res. 4:2070-2080 (2005).
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04665; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAA61237.1; ALT SEQ; mRNA.
DR EMBL; J04835; AAA61178.1; -; Genomic DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1Z44; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_G_Tsp_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWFC_1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS00184; VWFC_2; 1.
DR 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT-----
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT REGION 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 385 385
FT CARBOHYD 394 394
FT CARBOHYD 438 438
FT CARBOHYD 441 441
FT CARBOHYD 450 450
FT CARBOHYD 498 498
FT CARBOHYD 507 507
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT-----
FT Thrombospondin-1.
FT /FTID=PRO_0000035842.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...).
FT N-linked (GlcNAc...). (Potential).
FT C-linked (Man).
FT /FTID=CAR_000205.
FT O-linked (Fuc...).
FT /FTID=CAR_000206.
FT C-linked (Man).
FT /FTID=CAR_000207.
FT C-linked (Man).
FT /FTID=CAR_000208.
FT O-linked (Fuc...).
FT /FTID=CAR_000209.
FT C-linked (Man).
FT /FTID=CAR_000210.
FT O-linked (Fuc...).
FT /FTID=CAR_000211.
FT N-linked (GlcNAc...). (Potential).
FT N-linked (GlcNAc...).
FT Interchain (Probable).
FT Interchain (Probable).

Query Match 90.3%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNVRVVF 12

Db 208 FQGVNVRVVF 219

RESULT 5

TSPI_MOUSE STANDARD; PRT; 1170 AA.

AC P3541;

DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.

DT 01-JUN-1994, sequence version 1.

DT 07-MAR-2006, entry version 57.

DE Thrombospondin-1 precursor.

GN Name=Thb1; Synonyms=Tsp1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridea; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

[1]

RN NUCLEOTIDE SEQUENCE.

RX MEDLINE=92128941; PubMed=1774063;

RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,

RA Jenkins N.A.;

RA "Characterization of the murine thrombospondin gene.";

RT Genomics 11:587-600(1991).

[2]

RN NUCLEOTIDE SEQUENCE.

RX MEDLINE=92147683; PubMed=1371115;

RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,

RA Dixit V.M.;

RA "Characterization of mouse thrombospondin 2 sequence and expression

RT during cell growth and development.";

RT J. Biol. Chem. 267:3274-3281(1992).

[3]

RN NUCLEOTIDE SEQUENCE OF 1-490.

RX MEDLINE=90375546; PubMed=2398070;

RA Bornstein P., Alfi D., Devarayalu S., Framson P., Li P.;

RA "Characterization of the mouse thrombospondin gene and evaluation of

RT the role of the first intron in human gene expression.";

RT J. Biol. Chem. 265:16691-16698(1990).

[4]

RN PROTEIN SEQUENCE OF 19-37.

RP PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;

RA Chen H., Aeschlimann D., Nowlin J., Mosher D.F.;

RA "Expression and initial characterization of recombinant mouse

RT thrombospondin 1 and thrombospondin 3.";

RT FEBS Lett. 387:36-41(1996).

-!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and

CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,

CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-

CC V/beta-3 and alpha-1Ib/beta-3.

-!- SUBUNIT: Homotrimer; disulfide-linked.

-!- SIMILARITY: Belongs to the thrombospondin family.

-!- SIMILARITY: Contains 3 EGF-like domains.

-!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.

-!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

-!- SIMILARITY: Contains 3 TSP type-1 domains.

-!- SIMILARITY: Contains 7 TSP type-3 domains.

-!- SIMILARITY: Contains 1 WWFC domain.

CC

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC

EMBL; MG2470; AAA50611.1; -; Genomic DNA.

DR EMBL; MG2450; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2451; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2452; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2453; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2454; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2455; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2456; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2457; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2458; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2459; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; MG2460; AAA50611.1; JOINED; Genomic DNA.

```

FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
FT FT DISULFID 270 270 Interchain (Probable).
FT FT DISULFID 274 274 Interchain (Probable).
FT FT DISULFID 391 423 By similarity.
FT FT DISULFID 395 428 By similarity.
FT FT DISULFID 406 413 By similarity.
FT FT DISULFID 447 484 By similarity.
FT FT DISULFID 451 489 By similarity.
FT FT DISULFID 462 474 By similarity.
FT FT DISULFID 504 541 By similarity.
FT FT DISULFID 508 546 By similarity.
FT FT DISULFID 519 531 By similarity.
FT FT DISULFID 551 562 By similarity.
FT FT DISULFID 556 572 By similarity.
FT FT DISULFID 575 586 By similarity.
FT FT DISULFID 592 608 By similarity.
FT FT DISULFID 599 617 By similarity.
FT FT DISULFID 620 644 By similarity.
FT FT DISULFID 650 663 By similarity.
FT FT DISULFID 657 676 By similarity.
FT FT DISULFID 678 689 By similarity.
FT FT DISULFID 705 713 By similarity.
FT FT DISULFID 718 738 By similarity.
FT FT DISULFID 754 774 By similarity.
FT FT DISULFID 777 797 By similarity.
FT FT DISULFID 813 833 By similarity.
FT FT DISULFID 836 856 By similarity.
FT FT DISULFID 874 894 By similarity.
FT FT DISULFID 910 930 By similarity.
FT FT DISULFID 946 1167 By similarity.
FT FT CONFLICT 1025 1025 F -> L (in Ref. 2).
SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 90.3%; Score 56; DB 1; Length 1170;
Best local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLNVRVFV 12
Db 208 FQGVLLNVRVFV 219

RESULT 6
Q3TR40_MOUSE PRELIMINARY; PRT; 1170 AA.
AC Q3TR40;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 5.
DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DE clone: A530058N06 product: thrombospondin 1, full insert sequence.
GN Name=thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impiombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,

```

```

RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Putaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiuira K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RT (Genome Network Core Team) and the FANTOM Consortium;
RL "Antisense transcription in the Mammalian Transcriptome.";
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```


RT 60.770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Mateu Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaki E., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [8]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
 RA Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; AK163092; BAE37190.1; -; mRNA.
 DR MGI:98737; Thba1.
 DR GO:0005615; C:extracellular space; IDA.
 DR GO:0005615; C:extracellular space; RCA.
 DR GO:0016525; P:negative regulation of angiogenesis; IDA.
 DR InterPro; IPRO06210; EGF.
 DR InterPro; IPRO00742; EGF 3.
 DR InterPro; IPRO01881; EGF_Ca bd.
 DR InterPro; IPRO06209; EGF_like.
 DR InterPro; IPRO13032; EGF_like_reg.

DR InterPro; IPRO03129; Laminin_G_TSP_N.
 DR InterPro; IPRO00884; TSP1.
 DR InterPro; IPRO00885; TSP_1.
 DR InterPro; IPRO03367; tsg_3.
 DR InterPro; IPRO08859; TSP_C.
 DR InterPro; IPRO01007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 12.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; VMC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VMC; 1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; VWF_C; 1; 1.
 DR PROSITE; PS0184; VWF_C; 2; 1.
 SQ SEQUENCE 1170 AA; 129619 MW; 55BF04A20D91B194 CRC64;
 Query Match 90.3%; Score 56; DB 2; Length 1170;
 Best Local Similarity 91.7%; Pred. No. 0.22; Mismatches 0; Gaps 0;
 Matches 11; Conservative 0; Indels 1; Indels 0; Gaps 0;
 Oy 1 FQGVLNVRVVF 12
 Db 208 FQGVLNVRVVF 219
 RESULT 7
 Q71SA3 RAT PRELIMINARY; PRT; 1170 AA.
 ID Q71SA3
 AC Q71SA3
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 9.
 DE Thrombospondin 1.
 GN Name=Thp1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Sprague-Dawley;
 RA Iwabu A., Hirohata S., Kusachi S., Nakamura K., Murakami T.,
 RA Ninomiya Y., Tsuji T.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; AF309630; AAQ14549.1; -; mRNA.
 DR SMR; 834-1169.
 DR GO:0005576; C:extracellular region; IEA.
 DR GO:0005509; F:calcium ion binding; IEA.
 DR GO:0005198; F:structural molecule activity; IEA.
 DR GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPRO13320; ConA_like_subgrp.
 DR InterPro; IPRO06210; EGF.
 DR InterPro; IPRO00742; EGF 3.
 DR InterPro; IPRO01881; EGF_Ca bd.
 DR InterPro; IPRO06209; EGF_like.
 DR InterPro; IPRO13032; EGF_like_reg.
 DR InterPro; IPRO03129; Laminin_G_TSP_N.
 DR InterPro; IPRO00884; TSP1.
 DR InterPro; IPRO08085; TSP_1.
 DR InterPro; IPRO03367; tsg_3.
 DR InterPro; IPRO08859; TSP_C.

DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 12.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; VWC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; VWF_C; 1; UNKNOWN_1.
 DR PROSITE; PS00184; VWF_C; 2; 1.
 SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 90.3%; Score 56; DB 2; Length 1170;
 Best Local Similarity 91.7%; Pred. No. 0.22;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12
 ||||| |||||
 Db 208 FQGVLNVRVVF 219

RESULT 8

ID Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.
 AC Q80YQ1;
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2003, sequence version 1.
 DT 07-FEB-2006, entry version 17.
 DE Thrombospondin 1.
 GN Name=Thbs1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Director MGC Project;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted under the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; BC050917; AAH50917.1; -; mRNA.
 DR HSSP; P07996; 1LSL.
 DR SMR; Q80YQ1; 835-1170.
 DR Ensembl; ENSMUSG00000040152; Mus musculus.
 DR MGI; MGI:98737; Thbs1.
 DR GO; GO:0005615; C:extracellular space; RCA.
 DR GO; GO:0005615; C:extracellular space; IDA.
 DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
 DR InterPro; IPR013320; ConA_like_subgrp.
 DR InterPro; IPR006210; EGF.
 DR InterPro; IPR000742; EGF_3.
 DR InterPro; IPR001881; EGF_Ca_bd.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR013032; EGF-like reg.
 DR InterPro; IPR003129; Laminin_G_TSP_N.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; tsp_3.
 DR InterPro; IPR008859; TSP_C.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 12.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; VWC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; VWF_C; 1; UNKNOWN_1.
 DR PROSITE; PS00184; VWF_C; 2; 1.
 SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 90.3%; Score 56; DB 2; Length 1171;
 Best Local Similarity 91.7%; Pred. No. 0.22;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12
 ||||| |||||
 Db 208 FQGVLNVRVVF 219

RESULT 9

Q8CGB2 MOUSE
 ID Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.
 AC Q8CGB2;
 DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
 DT 01-MAR-2003, sequence version 1.
 DT 07-FEB-2006, entry version 22.
 DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-
 DE length enriched library, clone:G930018021 product:thrombospondin 1,
 DE full insert sequence).
 DE Name=Thbs1;
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor metastatized to lung. Tumor
 RC arose spontaneously;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Director MGC Project;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted under the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Trinchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinnwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C2CH II; TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RC Tissue submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6875(99)03004-9;
 RX Carninci P., Hayashizaki Y.;
 RA "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44 (1999).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K., Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Wilming L.G., Aidinis V., Allen J.E., Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L., Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M., Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R., Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G., di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G., Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M., Georgii-Heminger P., Gierantas T.R., Gojobori T., Green R.E., Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N., Hill D., Humnietek L., Iacono M., Ikeo K., Iwama A., Ishikawa T., Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H., Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K., Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J., Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L., Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K., Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P., Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O., Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G., Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M., Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y., Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y., Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B., Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K., Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A., Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K., Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C., Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J., Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y., Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T., Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N., Kawashima T., Kohjima M., Kondo S., Konno H., Nakano K., Ninomiya N., Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S., Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J., Hayashizaki Y.;
 RA "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563 (2005).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RG Riken Genome Exploration Research Group, and Genome Science Group (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566 (2005).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01286;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake E., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I., Miyazaki A., Sasaki K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S., Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [8]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [9]


```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 602 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 90.3%; Score 56; DB 1; Length 1173;
Best Local Similarity 91.7%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12
Db 211 FQGVLNVRVVF 222

RESULT 11
Q59E99 HUMAN
ID Q59E99_HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AB209912; BAD93149.1; -; mRNA.
CC SMR; Q59E99; 886-939, 889-1225.
CC Ensembl; ENSG00000137801; Homo sapiens.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:000515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR00742; EGF_3.
CC InterPro; IPR01881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR008084; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR003367; TSP_3.
CC InterPro; IPR008859; TSP_C.
```

```
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS0184; VWC_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16E57157B12 CRC64;

Query Match 90.3%; Score 56; DB 2; Length 1225;
Best Local Similarity 91.7%; Pred. No. 0.23;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLNVRVVF 12
Db 263 FQGVLNVRVVF 274

RESULT 12
Q5SPG5 BRARE
ID Q5SPG5_BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AL528866; CAI20599.1; -; Genomic_DNA.
CC SMR; Q5SPG5; 751-804, 754-1089.
CC Ensembl; ENSDARG0000010785; Danio rerio.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:000515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR002048; EF_hand_Ca_bd.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR00742; EGF_3.
CC InterPro; IPR01881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR008084; TSP1.
CC InterPro; IPR008085; TSP_1.
```

```

DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 2.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1090 AA; 5A9320504A22D836 CRC64;

Query Match 82.3%; Score 51; DB 2; Length 1090;
Best Local Similarity 83.3%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVNNVRVVF 12
Db 185 FMGVNQNVREVF 196

RESULT 13
Q4S758 TETNG
ID Q4S758 TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-FEB-2006, sequence version 1.
DE Chromosome 14 SCAF14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG0022976001;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Athouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC

```

Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License

```

CC EMBL; CAAE01014723; GAG03524.1; -; Genomic_DNA.
CC SMR; Q4S758; 811-1148.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; F:cell adhesion; IEA.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR013032; EGF-like reg.
CC InterPro; IPR003123; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR001007; VWFC_C.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS01186; EGF_2; UNKNOWN_1.
CC PROSITE; PS50026; EGF_3; 2.
CC PROSITE; PS50092; TSP1; 3.
CC PROSITE; PS01208; VWFC_1; 1.
CC PROSITE; PS50184; VWFC_2; 1.
KW Cell adhesion.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 82.3%; Score 51; DB 2; Length 1193;
Best Local Similarity 83.3%; Pred. No. 2;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVNNVRVVF 12
Db 187 FMGVNQNVREVF 198

RESULT 14
Q5U903 FIG
ID Q5U903 FIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang K., Mauc G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms  

Distributed under the Creative Commons Attribution-NoDerivs License


---



```

CC EMBL; AY773342; AAV38110.1; -; mRNA.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR001007; VWFC_C.

```


```

```
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SMO0209; TSP1; 1.
DR SMART; SMO0214; VWC; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
FT NON_TER 249
SQ SEQUENCE 249 AA; 27560 MW; 465D664BB0329C32 CRC64;

Query Match 80.6%; Score 50; DB 2; Length 249;
Best Local Similarity 90.9%; Pred.No. 0.58;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLVNVRVFP 12
Db 1 QGVLVNVRVFP 11

RESULT 15
Q4RLR5 TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, sequence version 1.
DE Chromosome 10 SCAF15019, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTNG0032374001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Luthalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC ENBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
CC EMBL; CAAE01015019; CAG10667.1; -; Genomic_DNA.
DR SMR; Q4RLR5; 834-887, 837-1171.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:haptan binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

```
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SMO0181; EGF; 2.
DR SMART; SMO0209; TSP1; 3.
DR SMART; SMO0210; TSPN; 1.
DR SMART; SMO0214; VWC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
KW Cell adhesion.
FT NON_TER 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 80.6%; Score 50; DB 2; Length 1171;
Best Local Similarity 83.3%; Pred.No. 3.1;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLVNVRVFP 12
Db 207 FQGVLVNVRVFP 218

Search completed: June 5, 2006, 22:42:36
Job time : 110.931 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-24

Perfect score: 62

Sequence: 1 FQGVNNRVFV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /EMC Celleria_SIDS3/ptodata/2/iaa/5 COMB.pap.*
 - 2: /EMC Celleria_SIDS3/ptodata/2/iaa/6 COMB.pap.*
 - 3: /EMC Celleria_SIDS3/ptodata/2/iaa/7 COMB.pap.*
 - 4: /EMC Celleria_SIDS3/ptodata/2/iaa/H COMB.pap.*
 - 5: /EMC Celleria_SIDS3/ptodata/2/iaa/PCITUS COMB.pap.*
 - 6: /EMC Celleria_SIDS3/ptodata/2/iaa/RE COMB.pap.*
 - 7: /EMC Celleria_SIDS3/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	90.3	825	2	US-09-949-002-482
2	56	90.3	831	2	US-09-939-853A-97
3	56	90.3	831	2	US-09-939-853A-98
4	56	90.3	1170	1	US-08-313-288B-20
5	56	90.3	1170	2	US-09-657-472-2
6	56	90.3	1170	2	US-09-949-002-350
7	38	61.3	1045	2	US-09-949-016-1112
8	38	61.3	1172	1	US-08-313-288B-19
9	38	61.3	1172	2	US-09-949-016-6333
10	37	59.7	740	2	US-09-134-000C-6441
11	37	59.7	943	2	US-09-583-110-4794
12	37	59.7	945	2	US-09-107-433-3340
13	36	58.1	61	2	US-09-270-767-61396
14	36	58.1	101	2	US-09-902-540-11494
15	36	58.1	175	2	US-09-252-991A-18826
16	36	58.1	302	2	US-09-248-796A-24789
17	36	58.1	304	2	US-09-270-767-45864
18	36	58.1	1288	2	US-08-762-428A-6
19	35	56.5	30	1	US-08-404-531B-37
20	35	56.5	30	2	US-08-476-900A-37
21	35	56.5	30	2	US-08-488-546A-37
22	35	56.5	65	2	US-09-540-236-3264
23	35	56.5	181	2	US-09-305-640-4
24	35	56.5	534	3	US-08-802-208B-4
25	35	56.5	640	1	US-08-671-978A-10
26	35	56.5	733	2	US-09-232-338-10

27	35	56.5	775	2	US-09-305-640-2	Sequence 2, Appli
28	35	56.5	775	2	US-09-948-429C-67	Sequence 67, Appl
29	35	56.5	994	2	US-09-543-681A-7288	Sequence 7288, Ap
30	35	56.5	1498	1	US-08-404-531B-28	Sequence 28, Appl
31	35	56.5	1498	1	US-08-404-531B-29	Sequence 29, Appl
32	35	56.5	1498	2	US-08-476-900A-28	Sequence 28, Appl
33	35	56.5	1498	2	US-08-476-900A-29	Sequence 29, Appl
34	35	56.5	1498	2	US-08-488-546A-28	Sequence 28, Appl
35	35	56.5	1498	2	US-08-488-546A-29	Sequence 29, Appl
36	35	56.5	1580	2	US-08-726-320-1	Sequence 1, Appli
37	35	56.5	1580	2	US-09-208-716-1	Sequence 1, Appli
38	35	56.5	1581	1	US-08-404-531B-6	Sequence 6, Appli
39	35	56.5	1581	2	US-08-476-900A-6	Sequence 6, Appli
40	35	56.5	1581	2	US-08-488-546A-6	Sequence 6, Appli
41	35	56.5	1581	2	US-08-726-320-3	Sequence 3, Appli
42	35	56.5	1581	2	US-08-726-320-4	Sequence 4, Appli
43	35	56.5	1581	2	US-09-208-716-3	Sequence 3, Appli
44	35	56.5	1581	2	US-09-208-716-4	Sequence 4, Appli
45	35	56.5	1582	1	US-08-404-531B-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL000790

; CURRENT APPLICATION NUMBER: US/09/949,002

; CURRENT FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/231,401

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 482

; LENGTH: 825

; TYPE: PRT

; ORGANISM: Human

US-09-949-002-482

Query Match 90.3% Score 56; DB 2; Length 825;

Best Local Similarity 91.7%; Pred.No. 0.043; 1; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVNNRVFV 12

|||||

Db 308 FQGVNNRVFV 319

RESULT 2

US-09-939-853A-97

; Sequence 97, Application US/09939853A

; Patent No. 6989232

; GENERAL INFORMATION:

; APPLICANT: Burgess et al.

; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-099

; CURRENT APPLICATION NUMBER: US/09/939,853A

; CURRENT FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: 60/228,191

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: 60/267,300

; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/269,961

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/277,337

; PRIOR FILING DATE: 2001-03-20

```
;
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          90.3%; Score 56; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.044;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY      1 FQGVLNVRVVF 12
      ||||| |||||
Db      208 FQGVLNVRVVF 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          90.3%; Score 56; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.044;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY      1 FQGVLNVRVVF 12
      ||||| |||||
Db      208 FQGVLNVRVVF 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          90.3%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLNVRVVF 12
      ||||| |||||
Db      208 FQGVLNVRVVF 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George O.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          90.3%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLNVRVVF 12
      ||||| |||||
Db      208 FQGVLNVRVVF 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-350

Query Match          90.3%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLNVRVFP 12
   ||||| |||||
Db 208 FQGVLNVRVFP 219

RESULT 7
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11112

Query Match          61.3%; Score 38; DB 2; Length 1045;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLNVRVFP 12
   ||:|:|:|:|
Db 281 FRGLQNVHLVF 292

RESULT 8
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
```

```
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match          61.3%; Score 38; DB 1; Length 1172;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLNVRVFP 12
   ||:|:|:|:|
Db 202 FRGLQNVHLVF 213

RESULT 9
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6333

Query Match          61.3%; Score 38; DB 2; Length 1172;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLNVRVFP 12
   ||:|:|:|:|
Db 202 FRGLQNVHLVF 213

RESULT 10
```

```
US-09-134-000C-6441
; Sequence 6441, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6441
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (11)..(11)
; OTHER INFORMATION: Amino acid 11 is Xaa wherein Xaa = any amino acid.
US-09-134-000C-6441
Query Match 59.7%; Score 37; DB 2; Length 740;
Best Local Similarity 41.7%; Pred. No. 1.3e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLLNVRVF 12
Db 728 FCGILGNINFIY 739

RESULT 11
US-09-583-110-4794
; Sequence 4794, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4794
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4794
Query Match 59.7%; Score 37; DB 2; Length 943;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLLNVR 9
Db 373 FEGVINNIK 381

RESULT 12
US-09-107-433-3340
; Sequence 3340, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

US-09-134-000C-6441
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3340:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 945 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...945
; SEQUENCE DESCRIPTION: SEQ ID NO: 3340:
US-09-107-433-3340
Query Match 59.7%; Score 37; DB 2; Length 945;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLLNVR 9
Db 375 FEGVINNIK 383

RESULT 13
US-09-270-767-61396
; Sequence 61396, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61396
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61396
Query Match 58.1%; Score 36; DB 2; Length 61;
```

Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQVLNNRVFV 12
Db 44 FFGVTVNRLFP 55

RESULT 14

US-09-902-540-11494
; Sequence 11494, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11494
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11494

Query Match 58.1%; Score 36; DB 2; Length 101;

Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLLNNRVF 10
Db 23 QGMLNRIRF 31

RESULT 15

US-09-252-991A-18826
; Sequence 18826, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18826
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18826

Query Match 58.1%; Score 36; DB 2; Length 175;

Best Local Similarity 63.6%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLLNNRVFV 12
Db 41 QGVLDVQGLF 51

Search completed: June 5, 2006, 22:48:52
Job time : 24.8966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78 6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-24

Perfect score: 62

Sequence: 1 FQGVNNRVFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	90.3	12	4	US-10-474-213-28
2	56	90.3	240	4	Sequence 28, Appl
3	56	90.3	240	5	Sequence 40, Appl
4	56	90.3	432	5	Sequence 1020, Ap
5	56	90.3	432	5	Sequence 1022, Ap
6	56	90.3	459	6	Sequence 462, App
7	56	90.3	466	3	Sequence 1047, Ap
8	56	90.3	555	6	Sequence 454, App
9	56	90.3	578	6	Sequence 456, App
10	56	90.3	685	6	Sequence 452, App
11	56	90.3	804	6	Sequence 453, App
12	56	90.3	828	6	Sequence 455, App
13	56	90.3	831	3	Sequence 97, Appl
14	56	90.3	831	3	Sequence 98, Appl
15	56	90.3	855	6	Sequence 461, App
16	56	90.3	1000	6	Sequence 457, App
17	56	90.3	1105	6	Sequence 458, App
18	56	90.3	1150	4	Sequence 1, Appli
19	56	90.3	1152	3	Sequence 1, Appli
20	56	90.3	1169	5	Sequence 7, Appli
21	56	90.3	1170	4	Sequence 12, Appl
22	56	90.3	1170	4	Sequence 2, Appli
23	56	90.3	1170	4	Sequence 114, App
24	56	90.3	1170	4	Sequence 2, Appli
25	56	90.3	1170	4	Sequence 1170, Ap
26	56	90.3	1170	4	Sequence 38, Appl
27	56	90.3	1170	4	Sequence 482, App

28 56 90.3 1170 4 US-10-419-462-38 Sequence 38, Appl
29 56 90.3 1170 5 US-10-741-600-1018 Sequence 1018, Ap
30 56 90.3 1170 5 US-10-741-600-1019 Sequence 1019, Ap
31 56 90.3 1170 5 US-10-741-600-1021 Sequence 1021, Ap
32 56 90.3 1170 5 US-10-782-968-38 Sequence 38, Appl
33 56 90.3 1170 5 US-10-849-989-44 Sequence 44, Appl
34 56 90.3 1170 5 US-10-631-467-548 Sequence 548, App
35 56 90.3 1170 5 US-10-631-467-1376 Sequence 1376, Ap
36 56 90.3 1170 5 US-10-831-997-2 Sequence 2, Appli
37 56 90.3 1170 5 US-10-995-561-594 Sequence 594, App
38 56 90.3 1170 5 US-10-995-561-595 Sequence 595, App
39 56 90.3 1170 5 US-10-995-561-596 Sequence 596, App
40 56 90.3 1170 6 US-11-037-713-51 Sequence 51, Appl
41 56 90.3 1170 6 US-11-046-644-28 Sequence 28, Appl
42 56 90.3 1170 6 US-11-046-456-28 Sequence 28, Appl
43 40 64.5 15 4 US-10-285-394-153 Sequence 153, App
44 40 64.5 597 4 US-10-282-122A-76883 Sequence 76883, A
45 39 62.9 226 5 US-10-467-657-2428 Sequence 2428, Ap

ALIGNMENTS

RESULT 1
US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Krutzsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/10535
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 90.3%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFVF 12

Db 1 FQGVNNRVFVF 12

RESULT 2

US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J.
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40

; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match 90.3%; Score 56; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNVRFV 12
| | | | | | | | | |
Db 190 FQGVQNVRVF 201

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match 90.3%; Score 56; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNVRFV 12
| | | | | | | | | |
Db 190 FQGVQNVRVF 201

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 90.3%; Score 56; DB 5; Length 432;

Best Local Similarity 91.7%; Pred. No. 0.069;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNVRFV 12
| | | | | | | | | |
Db 208 FQGVQNVRVF 219

RESULT 5
US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 90.3%; Score 56; DB 5; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.069;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNVRFV 12
| | | | | | | | | |
Db 208 FQGVQNVRVF 219

RESULT 6
US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match 90.3%; Score 56; DB 6; Length 459;
Best Local Similarity 91.7%; Pred. No. 0.074;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNVRFV 12
| | | | | | | | | |
Db 208 FQGVQNVRVF 219

RESULT 7
US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925.301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 90.3%; Score 56; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.075;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFV 12
Db 261 FQGVQNVRFV 272

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 90.3%; Score 56; DB 6; Length 555;
Best Local Similarity 91.7%; Pred. No. 0.091;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFV 12
Db 208 FQGVQNVRFV 219

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 90.3%; Score 56; DB 6; Length 578;
Best Local Similarity 91.7%; Pred. No. 0.095;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFV 12
Db 208 FQGVQNVRFV 219

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 90.3%; Score 56; DB 6; Length 685;
Best Local Similarity 91.7%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFV 12
Db 208 FQGVQNVRFV 219

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 90.3%; Score 56; DB 6; Length 804;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFV 12
Db 208 FQGVQNVRFV 219

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```

; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match          90.3%; Score 56; DB 6; Length 828;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFV 12
Db      208 FQGVLLNNRVFV 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR FILING DATE: 2001-08-27
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 2001-02-08
; PRIOR FILING DATE: 2001-02-08
; PRIOR FILING DATE: 2001-02-20
; PRIOR FILING DATE: 2001-02-20
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          90.3%; Score 56; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFV 12
Db      208 FQGVLLNNRVFV 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR FILING DATE: 2001-08-27
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 2001-02-08
; PRIOR FILING DATE: 2001-02-08
; PRIOR FILING DATE: 2001-02-20
; PRIOR FILING DATE: 2001-02-20
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus

```

```

US-09-939-853A-98

Query Match          90.3%; Score 56; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFV 12
Db      208 FQGVLLNNRVFV 219

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match          90.3%; Score 56; DB 6; Length 855;
Best Local Similarity 91.7%; Pred. No. 0.15;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFV 12
Db      208 FQGVLLNNRVFV 219

Search completed: June 6, 2006, 00:00:09
Job time : 79.6207 secs

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-24

Perfect score: 62

Sequence: 1 FQGVNNRVFV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

1: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US09_NEW_PUB.pap:*
2: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US06_NEW_PUB.pap:*
3: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US07_NEW_PUB.pap:*
4: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US08_NEW_PUB.pap:*
5: /EMC_Celerra_SIDS3/ptodata/1/pubppa/PCT_NEW_PUB.pap:*
6: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US10_NEW_PUB.pap:*
7: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US11_NEW_PUB.pap:*
8: /EMC_Celerra_SIDS3/ptodata/1/pubppa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	58.1	132	6	US-10-953-349-39312
2	34	54.8	387	7	US-11-293-697-3087
3	34	54.8	1504	6	US-10-505-928-662
4	33	53.2	164	6	US-10-953-349-1159
5	33	53.2	249	6	US-10-953-349-1158
6	33	53.2	250	6	US-10-953-349-1157
7	33	53.2	262	6	US-10-953-349-24927
8	33	53.2	310	6	US-10-953-349-24926
9	33	53.2	315	6	US-10-953-349-9204
10	33	53.2	331	6	US-10-953-349-34674
11	33	53.2	358	6	US-10-953-349-34673
12	33	53.2	381	6	US-10-953-349-24925
13	33	53.2	382	6	US-10-953-349-34672
14	33	53.2	400	6	US-10-953-349-9203
15	33	53.2	401	6	US-10-953-349-9202
16	33	53.2	843	6	US-10-953-349-16527
17	33	53.2	962	6	US-10-953-349-16526
18	33	53.2	1014	6	US-10-953-349-16525
19	32.5	52.4	101	6	US-10-953-349-2569
20	32	51.6	182	6	US-10-953-349-15318
21	32	51.6	297	7	US-11-293-697-4132
22	32	51.6	353	6	US-10-953-349-21136
23	32	51.6	447	6	US-10-953-349-21135
24	32	51.6	470	6	US-10-953-349-21134
25	31	50.0	210	6	US-10-953-349-21828

26	31	50.0	223	6	US-10-953-349-21827	Sequence 21827, A
27	31	50.0	273	6	US-10-953-349-21826	Sequence 21826, A
28	31	50.0	368	6	US-10-953-349-20148	Sequence 20148, A
29	31	50.0	407	6	US-10-953-349-26200	Sequence 26200, A
30	31	50.0	437	6	US-10-953-349-20147	Sequence 20147, A
31	31	50.0	444	6	US-10-953-349-20146	Sequence 20146, A
32	31	50.0	524	7	US-11-293-697-3442	Sequence 3442, Ap
33	31	50.0	980	7	US-11-242-505A-36	Sequence 36, Appl
34	30	48.4	397	6	US-10-953-349-23556	Sequence 23556, A
35	30	48.4	437	6	US-10-504-120-32	Sequence 32, Appl
36	30	48.4	440	6	US-10-953-349-1348	Sequence 1348, Ap
37	30	48.4	525	6	US-10-953-349-13207	Sequence 13207, A
38	30	48.4	528	6	US-10-953-349-23555	Sequence 23555, A
39	30	48.4	531	6	US-10-953-349-13004	Sequence 13004, A
40	30	48.4	538	6	US-10-953-349-13003	Sequence 13003, A
41	30	48.4	543	6	US-10-953-349-23554	Sequence 23554, A
42	30	48.4	546	6	US-10-953-349-13002	Sequence 13002, A
43	30	48.4	560	6	US-10-953-349-31206	Sequence 31206, A
44	30	48.4	627	7	US-11-249-111-88	Sequence 88, Appl
45	29	46.8	108	7	US-11-293-697-4588	Sequence 4588, Ap

ALIGNMENTS

RESULT 1

US-10-953-349-39312
; Sequence 39312, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39312
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-39312

Query Match 58.1%; Score 36; DB 6; Length 132;
Best Local Similarity 54.5%; Pred. No. 3.1;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVNNRVFV 11
Db 66 FEGILNNVFI 76

RESULT 2

US-11-293-697-3087
; Sequence 3087, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293.697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108.260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3087
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3087


```
Db      133 EGVSNKIRFI 142
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match      53.2%; Score 33; DB 6; Length 331;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      2 QGVLLNNRVFV 11
Db      225 QGVLFNIQYV 234
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

RESULT 11
US-10-953-349-34673
; Sequence 34673, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

Query Match      53.2%; Score 33; DB 6; Length 310;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      2 QGVLLNNRVFV 11
Db      181 EGVSNKIRFI 190
; SEQ ID NO 9204
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9204

RESULT 9
US-10-953-349-9204
; Sequence 9204, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9204
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9204

Query Match      53.2%; Score 33; DB 6; Length 315;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 FQGVLLNNRVFV 12
Db      82 FEAVVDRVRLVF 93
; SEQ ID NO 24925
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-24925

RESULT 10
US-10-953-349-34674
; Sequence 34674, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match      53.2%; Score 33; DB 6; Length 381;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      2 QGVLLNNRVFV 11
Db      252 EGVSNKIRFI 261
; SEQ ID NO 24925
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-24925

RESULT 13
US-10-953-349-34672
```

```
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match      53.2%; Score 33; DB 6; Length 331;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      2 QGVLLNNRVFV 11
Db      225 QGVLFNIQYV 234
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

RESULT 11
US-10-953-349-34673
; Sequence 34673, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

Query Match      53.2%; Score 33; DB 6; Length 358;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      2 QGVLLNNRVFV 11
Db      252 QGVLFNIQYV 261
; SEQ ID NO 24925
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-24925

RESULT 12
US-10-953-349-24925
; Sequence 24925, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24925
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-24925

Query Match      53.2%; Score 33; DB 6; Length 381;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      2 QGVLLNNRVFV 11
Db      252 EGVSNKIRFI 261
; SEQ ID NO 24925
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-24925

RESULT 13
US-10-953-349-34672
```

```
; Sequence 34672, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34672
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34672

Query Match      53.2%; Score 33; DB 6; Length 382;
Best Local Similarity 60.0%; Pred. No. 38;
Matches      6; Conservative      3; Mismatches      1; Indels      0; Gaps      0;

Qy      2 QGVLLNNRVFV 11
Db      276 QGVLFNIQYV 285

RESULT 14
US-10-953-349-9203
; Sequence 9203, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9203
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9203

Query Match      53.2%; Score 33; DB 6; Length 400;
Best Local Similarity 50.0%; Pred. No. 40;
Matches      6; Conservative      3; Mismatches      3; Indels      0; Gaps      0;

Qy      1 FQGVLLNNRVFV 12
Db      167 FEAVVDRVRLVF 178

RESULT 15
US-10-953-349-9202
; Sequence 9202, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9202
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9202
```

```
Query Match      53.2%; Score 33; DB 6; Length 401;
Best Local Similarity 50.0%; Pred. No. 40;
Matches      6; Conservative      3; Mismatches      3; Indels      0; Gaps      0;

Qy      1 FQGVLLNNRVFV 12
Db      168 FEAVVDRVRLVF 179

Search completed: June 6, 2006, 00:12:55
Job time : 3.82414 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-25
Perfect score: 59
Sequence: 1 AQQVLQNVRFVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep 8: *
1: Genesep1980s: *
2: Genesep1990s: *
3: Genesep2000s: *
4: Genesep2001s: *
5: Genesep2002s: *
6: Genesep2003as: *
7: Genesep2003bs: *
8: Genesep2004s: *
9: Genesep2005s: *
10: Genesep2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	12	4	AAB35370 Alpha3bet
2	55	93.2	11	4	AAB35357 Alpha3bet
3	55	93.2	12	4	AAB35352 Alpha3bet
4	55	93.2	12	4	AAB35378 Alpha3bet
5	55	93.2	12	6	ABG72834 Thrombos
6	55	93.2	240	8	ADL70641 Human thr
7	55	93.2	432	8	ADQ39359 Human thr
8	55	93.2	432	8	ADQ39357 Human myo
9	55	93.2	459	4	AAU02916 Angiotens
10	55	93.2	466	3	ABA43602 Human can
11	55	93.2	546	4	AAU02915 Angiotens
12	55	93.2	548	7	ADN02474 TSF polyp
13	55	93.2	555	4	AAU02914 Angiotens
14	55	93.2	731	4	AAU02913 Angiotens
15	55	93.2	1152	3	AAB00042 Human thr
16	55	93.2	1152	5	AAU74771 Human thr
17	55	93.2	1152	5	ABE82285 Human thr
18	55	93.2	1170	4	AAB74450 Human var
19	55	93.2	1170	4	AAB90800 Human she
20	55	93.2	1170	5	AAB25030 Human thr
21	55	93.2	1170	5	AAU75315 Human thr
22	55	93.2	1170	6	ABP96780 Human COP
23	55	93.2	1170	6	ABU03474 Angiogene

24	55	93.2	1170	6	ABG74673 Human THB
25	55	93.2	1170	6	AAE36228 Human THB
26	55	93.2	1170	7	ABR62059 Human thr
27	55	93.2	1170	7	ADN39852 Cancer/van
28	55	93.2	1170	8	ADJ76124 Marker ge
29	55	93.2	1170	8	ADJ75296 Marker ge
30	55	93.2	1170	8	ADL70639 Human thr
31	55	93.2	1170	8	ADL35874 Human thr
32	55	93.2	1170	8	ADQ26070 Thrombos
33	55	93.2	1170	8	ADP54179 Human PRO
34	55	93.2	1170	8	ADQ39358 Human myo
35	55	93.2	1170	8	ADQ39356 Human myo
36	55	93.2	1170	8	ADQ39355 Human myo
37	55	93.2	1170	9	ADZ21688 Thrombos
38	55	93.2	1170	9	ABE87781 Human thr
39	55	93.2	1170	9	ABE46751 Human thr
40	52	88.1	12	4	AAB35373 Alpha3bet
41	52	88.1	12	4	AAB35381 Alpha3bet
42	51	86.4	12	4	AAB35364 Alpha3bet
43	51	86.4	12	4	AAB35374 Alpha3bet
44	50	84.7	12	4	AAB35371 Alpha3bet
45	50	84.7	12	4	AAB35368 Alpha3bet

ALIGNMENTS

RESULT 1
AAB35370
ID AAB35370 standard; peptide; 12 AA.
AC AAB35370;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3bet1 integrin binding peptide #35.
XX
KW Alpha3bet1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
XX Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3bet1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
XX of the invention
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 59; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQGVQLQNVRFVF 12
 Db 1 AQGVQLQNVRFVF 12

RESULT 2

AAB35357
 ID AAB35357 standard; peptide; 11 AA.

XX
 AC AAB35357;

XX
 DT 08-MAY-2001 (first entry)

XX
 DE Alpha3betal integrin binding peptide #22.

XX KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
 KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX DR WPI; 2001-182656/18.

XX PT New peptides that bind to or are recognized by alpha3-betal integrins,
 PT useful for inhibiting cell adhesion to extracellular matrix, cell
 PT motility and proliferation and for treating rheumatoid arthritis and
 PT cancer.

XX PS Claim 4; Page 34; 84pp; English.

XX CC The present invention provides a number of peptides which bind to
 CC alpha3betal integrins. They are useful in the modulation of cell adhesion
 CC and motility, and in the treatment of cancer, diabetic retinopathy,
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
 CC and restenosis. The present sequence is an example of one of the peptides
 CC of the invention

XX SQ Sequence 11 AA;

Query Match 93.2%; Score 55; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.00083;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
 Db 1 QGVQLQNVRFVF 11

RESULT 3

AAB35352
 ID AAB35352 standard; peptide; 12 AA.

XX
 AC AAB35352;

XX
 DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #17.

XX

KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
 KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX DR WPI; 2001-182656/18.

XX PT New peptides that bind to or are recognized by alpha3-betal integrins,
 PT useful for inhibiting cell adhesion to extracellular matrix, cell
 PT motility and proliferation and for treating rheumatoid arthritis and
 PT cancer.

XX PS Claim 4; Page 34; 84pp; English.

XX CC The present invention provides a number of peptides which bind to
 CC alpha3betal integrins. They are useful in the modulation of cell adhesion
 CC and motility, and in the treatment of cancer, diabetic retinopathy,
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
 CC and restenosis. The present sequence is an example of one of the peptides
 CC of the invention

XX SQ Sequence 12 AA;

Query Match 93.2%; Score 55; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00092;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
 Db 2 QGVQLQNVRFVF 12

RESULT 4

AAB35378
 ID AAB35378 standard; peptide; 12 AA.

XX
 AC AAB35378;

XX DT 08-MAY-2001 (first entry)

XX DE Alpha3betal integrin binding peptide #43.

XX KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
 KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
 KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX PN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Roberts DD, Kruttsch HC;

XX WPI; 2001-182656/18.
 XX New peptides that bind to or are recognized by alpha3-beta1 integrins,
 PT useful for inhibiting cell adhesion to extracellular matrix, cell
 PT motility and proliferation and for treating rheumatoid arthritis and
 PT cancer.
 XX Example 2; Page 34; 84pp; English.
 XX The present invention provides a number of peptides which bind to
 CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
 CC and motility, and in the treatment of cancer, diabetic retinopathy,
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
 CC and restenosis. The present sequence is an example of one of the peptides
 CC of the invention
 XX Sequence 12 AA;
 SQ

Query Match 93.2%; Score 55; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00092;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 QGVQLQVRVFV 12
 Db 2 QGVQLQVRVFV 12
 |||||

RESULT 5
 ABG72834
 ID ABG72834 standard; peptide; 12 AA.
 XX
 AC ABG72834;
 XX
 DT 24-FEB-2003 (first entry)
 XX
 DE Thrombospondin-1 sequence containing synthetic peptide.
 XX
 KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
 KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
 KW colon cancer; small-cell lung cancer; SCLC; melanoma.
 XX
 OS Synthetic.
 XX
 FN WO200281630-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 03-APR-2002; 2002WO-US010535.
 XX
 PR 06-APR-2001; 2001US-0281994P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Roberts DD, Krutzsch HC;
 XX
 DR WPI; 2003-103329/09.
 XX
 PT A new diagnosis for cancer other than prostate cancer in a mammal useful
 PT to detect cancer including lung cancer, particularly small cell lung
 PT cancer and melanoma comprises detecting semenogelin in a sample.
 XX
 PS Example 1; Page 14; 32pp; English.
 XX
 CC The invention relates to diagnosing cancer other than prostate cancer in
 CC a male mammal, comprising assaying a test sample for increased level of
 CC semenogelin, or cancer in a female by assaying for the presence of
 CC semenogelin. Administering a semenogelin protein or polypeptide fragment
 CC or a semenogelin-specific antibody or active fragment, or a recombinant
 CC vector expressing the protein or antibody, is useful for inducing an
 CC immune response to a cancer in a mammal, where the cancer is not prostate
 CC cancer and semenogelin is a marker. The invention is used to diagnose
 CC cancer, particularly of epithelial origin such as lung cancer, papillary

CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
 CC (SCLC), or a melanoma. The present sequence represents the amino acid
 CC sequence of the thrombospondin-1 sequence containing synthetic peptide
 XX which binds to alpha-3-beta-1 integrin
 SQ Sequence 12 AA;
 Query Match 93.2%; Score 55; DB 6; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.00092;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 QGVQLQVRVFV 12
 Db 2 QGVQLQVRVFV 12
 |||||

RESULT 6
 ADL70641
 ID ADL70641 standard; protein; 240 AA.
 XX
 AC ADL70641;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human thrombospondin-1 N-terminal domain.
 XX
 KW Human; thrombospondin-1; epitope; cancer; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 23..32
 FT /note= "Heparin binding region"
 FT Region 77..82
 FT /note= "Heparin binding region"
 FT Region 151..164
 FT /note= "Fibrinogen binding region"
 XX
 FN WO2004018995-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2003; 2003WO-US026023.
 XX
 PR 23-AUG-2002; 2002US-0405494P.
 PR 21-APR-2003; 2003US-00419462.
 XX
 PA (WILL/) WILLIAMS K J.
 XX
 PI Williams KJ;
 XX
 DR WPI; 2004-226901/21.
 XX
 PT New purified thrombospondin fragment extracted from a body fluid, useful
 PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
 PT or leukemia or as calibrators, indicators, immunogens and analytes.
 XX
 PS Disclosure; SEQ ID NO 40; 76pp; English.
 XX
 CC The present sequence is that of the N-terminal domain of human
 CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
 CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
 CC clinical assays for cancer and for generation of antibodies and other
 CC binding agents. A method that distinguishes TSP from a TSP fragment or
 CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
 CC or portion as a target for a binding molecule, e.g. an antibody, to
 CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
 CC epitope present in TSP but not in the fragment or portion to obtain a
 CC quantitation of TSP only; and (3) using the difference between (1) and
 CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
 CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
 CC the TSP fragment or portion is performed in order to detect the presence,
 CC or monitor the course, of a disease or condition selected from cancer,

CC renal failure, renal disease, atopic dermatitis, vasculitis, acute
 CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
 CC infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
 CC nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
 CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
 CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
 CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition
 CC associated with clotting, a condition associated with platelet
 CC activation, a condition associated with intravascular platelet
 CC activation, a condition associated with consumption of platelets, heparin
 CC -induced thrombocytopenia, disseminated intravascular coagulation,
 CC intravascular coagulation, extravascular coagulation, a condition
 CC associated with endothelial activation, a condition associated with
 CC production and/or release of thrombospondin and/or a thrombospondin
 CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 CC reaction, an aspartame reaction, atopic dermatitis, eczema,
 CC hypersensitivity, scleroderma, conditions associated with plugging of
 CC vessels, a condition associated with a cryofibrinogen, a condition
 CC associated with a cryoglobulin, and a condition associated with an anti-
 CC cardiolipin antibody. The cancer is selected from adenoma,
 CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system,
 CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
 CC biliary system, pancreas, head, neck, endocrine system, reproductive
 CC system (male or female), genitourinary system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.

XX SQ Sequence 240 AA;

Query Match 93.2%; Score 55; DB 8; Length 240;

Best Local Similarity 100.0%; Pred. No. 0.027;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQNRVRFV 12

DB 191 QGVQLQNRVRFV 201

RESULT 7

ID ADQ39359

XX ADQ39359 standard; protein; 432 AA.

AC ADQ39359;

DT 18-NOV-2004 (first entry)

DE Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW cardiant; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

PN 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

PR 10-MAR-2003; 2003US-0453135P.

PR 30-APR-2003; 2003US-0466412P.

PR 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

PA Cargill M, Devlin JJ, Iakubova O;

XX Cargill M, Devlin JJ, Iakubova O;

PI Cargill M, Devlin JJ, Iakubova O;

XX Cargill M, Devlin JJ, Iakubova O;

DR WPI; 2004-533949/51.

XX N-PSDB; ADQ38531.

PT Identifying an individual who has an altered risk for developing

PT myocardial infarction by detecting a single nucleotide polymorphism in

PT the individual's nucleic acids.

XX Claim 10; SEQ ID NO 1022; 145pp; English.

CC The invention relates to a novel method for identifying an individual who

CC has an altered risk for developing myocardial infarction. The method

CC comprises detecting a single nucleotide polymorphism (SNP) in any one of

CC the nucleotide sequences given in the specification in the individual's

CC nucleic acids, where the presence of the SNP is correlated with an

CC altered risk for myocardial infarction in the individual. The invention

CC further comprises: an isolated nucleic acid molecule comprising at least

CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in

CC the specification or its complement and encoding any one of the amino

CC acid sequences given in the specification; an isolated polypeptide

CC comprising an amino acid sequence given in the specification; an antibody

CC that specifically binds to the polypeptide or its antigen-binding

CC fragment; an amplified polynucleotide containing an SNP given in the

CC specification and which is between about 16 and 1000 nucleotides in

CC length; a kit for detecting an SNP in a nucleic acid, comprising the

CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a

CC nucleic acid molecule; a method of detecting a variant polypeptide; and a

CC method for identifying an agent useful in treating or preventing

CC myocardial infarction. The novel detection method has cardiant activity.

CC The nucleic acids of the invention may be used in gene therapy. The

CC method is useful in identifying an individual who has an increased or

CC decreased risk for developing myocardial infarction and for preparing a

CC composition for treating or preventing myocardial infarction. This

CC sequence represents the protein of a human myocardial infarction-

CC associated gene containing one or more SNPs of the invention. Note: This

CC sequence was not shown in the specification. The sequence has come from

CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 432 AA;

Query Match 93.2%; Score 55; DB 8; Length 432;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQNRVRFV 12

DB 209 QGVQLQNRVRFV 219

RESULT 8

ID ADQ39357

XX ADQ39357 standard; protein; 432 AA.

AC ADQ39357;

XX 18-NOV-2004 (first entry)

DT Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

DE Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX cardiant; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

PN 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

PR 10-MAR-2003; 2003US-0453135P.

PR 30-APR-2003; 2003US-0466412P.

PR 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX N-PSDB; ADQ38529.
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX Claim 10; SEQ ID NO 1020; 145pp; English.
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNPs of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
XX
SQ Sequence 432 AA;
Query Match 93.2%; Score 55; DB 8; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 QGVQLQNVRFVF 12
Db 209 QGVQLQNVRFVF 219
RESULT 9
AAU02916
ID AAU02916 standard; protein; 459 AA.
XX
XX AAU02916;
XX
XX 12-SEP-2001 (first entry)
XX
XX Angiotensin converting enzyme (ACEV) splice variant protein #16.
XX
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX platelet-derived endothelial cell growth factor; cardiovascular disease;
XX cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX myocardial infarction; coronary arterial thrombosis; renal disease;
XX diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
XX multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
XX vascular disorder; asbestosis.

XX OS Homo sapiens.
XX WO200136632-A2.
XX PD 25-MAY-2001.
XX PF 17-NOV-2000; 2000WO-IL000766.
XX PR 17-NOV-1999; 99IL-00132978.
XX PR 10-DEC-1999; 99IL-00133455.
XX PA (COMP-) COMPUGEN LTD.
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX N-PSDB; AAS06016.
XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX Claim 4; Fig 16; 519pp; English.
XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
XX
SQ Sequence 459 AA;
Query Match 93.2%; Score 55; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 QGVQLQNVRFVF 12
Db 209 QGVQLQNVRFVF 219
RESULT 10
AAB43602
ID AAB43602 standard; protein; 466 AA.
XX
XX AAB43602;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated protein sequence SEQ ID NO:1047.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
XX antidiabetic; antiastrumatic; antirheumatic; antiarthritic; antiviral;
XX antiinflammatory; antichyroid; antiallergic; antibacterial; cardiac;
XX dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
XX vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening.

XX Homo sapiens.
 OS
 KW WO20005350-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US005882.
 PF
 XX 12-MAR-1999; 99US-0124270P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587533/55.
 DR
 XX N-PSDB; AAC77811.
 DR
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 PT
 XX Claim 11; Page 1636-1638; 2352pp; English.
 PS
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB44398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnarary; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nontropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 466 AA;

Query Match 93.2%; Score 55; DB 3; Length 466;
 Best Local Similarity 100.0%; Pred. No. 0.058;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
 |||||
 Db 262 QGVQLQNVRFVF 272

RESULT 11
 AAU02915
 ID AAU02915 standard; protein; 546 AA.
 XX
 AC AAU02915;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200136632-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL000766.
 PR
 XX 17-NOV-1999; 99IL-00132978.
 PR
 XX 10-DEC-1999; 99IL-00133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI
 XX WPI; 2001-336004/35.
 DR
 XX N-PSDB; AAS06015.
 DR
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 XX Claim 4; Fig 15; 519pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 546 AA;

Query Match 93.2%; Score 55; DB 4; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.069;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
 |||||
 Db 209 QGVQLQNVRFVF 219

RESULT 12
 ADN02474
 ID ADN02474 standard; protein; 548 AA.
 XX
 AC ADN02474;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE TSF polypeptide.
 XX
 KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
 XX
 OS Homo sapiens.
 XX
 PN CN1401387-A.
 XX
 PD 12-MAR-2003.

XX PF 21-AUG-2002; 2002CN-00129408.
 XX PR 21-AUG-2002; 2002CN-00129408.
 XX PA (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
 XX PI Han Z, Liu P;
 XX DR WPI; 2003-469302/45.
 XX DR N-PSDB; ADN02475.
 XX PT Tumor suppressing polypeptide TSF and gene therapy vector composition.
 XX PS Claim 2; SEQ ID NO 1; 13pp; Chinese.
 XX CC The present invention relates to a novel recombinant adenovirus vector mediated anti-neoplastic composition is prepared through cloning the cDNA sequence from the human peripheral blood cell by specific primer and reverse transcription-polymerase chain reaction (RT-PCR) method for coding TSF polypeptide, construction in human embryonic kidney cell 293 by AdEasy system, and packaging and expressing the recombinant adenovirus CC vector of TSF. It can suppress the growth and transfer of cancer. The CC present sequence represents the TSF polypeptide.
 XX SQ Sequence 548 AA;
 Query Match 93.2%; Score 55; DB 7; Length 548;
 Best Local Similarity 100.0%; Pred. No. 0.069;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 QGVQLQNVRFVF 12
 |||||
 DB 209 QGVQLQNVRFVF 219
 RESULT 13
 AAU02914
 ID AAU02914 standard; protein; 555 AA.
 XX AC AAU02914;
 XX DT 12-SEP-2001 (first entry)
 XX DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
 XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX OS Homo sapiens.
 XX PN WO200136632-A2.
 XX PD 25-MAY-2001.
 XX PF 17-NOV-2000; 2000WO-IL000766.
 XX PR 17-NOV-1999; 99IL-00132978.
 XX PR 10-DEC-1999; 99IL-00133455.
 XX PA (COMP-) COMPUGEN LTD.
 XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX WPI; 2001-336004/35.
 DR

DR N-PSDB; AAS06014.
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX Claim 4; Fig 14; 519pp; English.
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding to the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX SQ Sequence 555 AA;
 Query Match 93.2%; Score 55; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 QGVQLQNVRFVF 12
 |||||
 DB 209 QGVQLQNVRFVF 219
 RESULT 14
 AAU02913
 ID AAU02913 standard; protein; 731 AA.
 XX AC AAU02913;
 XX DT 12-SEP-2001 (first entry)
 XX DE Angiotensin converting enzyme (ACEV) splice variant protein #13.
 XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX OS Homo sapiens.
 XX PN WO200136632-A2.
 XX PD 25-MAY-2001.
 XX PF 17-NOV-2000; 2000WO-IL000766.
 XX PR 17-NOV-1999; 99IL-00132978.
 XX PR 10-DEC-1999; 99IL-00133455.
 XX PA (COMP-) COMPUGEN LTD.
 XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX WPI; 2001-336004/35.
 DR

DR N-PSDB; AAS06013.
XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 13; 519pp; English.
XX
CC The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variant of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 731 AA;
Query Match 93.2%; Score 55; DB 4; Length 731;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 QGVQLQNVRFVF 12
Db 209 QGVQLQNVRFVF 219
RESULT 15
AAB00042
ID AAB00042 standard; protein; 1152 AA.
XX
AC AAB00042;
XX
DT 08-NOV-2000 (first entry)
XX
DE Human thrombospondon-1 (TSP-1).
XX
KW TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
KW diabetic retinopathy; corneal graft rejection; glaucoma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 361..416
FT /label= Type 1 repeat region
FT Region 417..473
FT /label= Type 1 repeat region
FT Region 474..530
FT /label= Type 1 repeat region
XX
PN WO20004908-A2.
XX
PD 03-AUG-2000.
XX
PF 01-FEB-2000; 2000WO-US002482.
XX
PR 01-FEB-1999; 99US-0118053P.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Lawler JW;
XX
DR WPI; 2000-514823/46.

XX Nucleic acids encoding chimeric proteins such as cartilage oligomeric
PT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
PT inhibiting angiogenesis and treating diseases such as cancer.
XX
PS Disclosure; Fig 1; 40pp; English.
XX
CC New nucleic acids are described which encode a protein comprising the
CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
CC the TGF (transforming growth factor)-beta activation region of human TSP-
CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
CC tumours in mice models. Thus the nucleic acids and proteins may be useful
CC for treating angiogenesis related diseases such as cancer (by reducing
CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
CC for treating human immunodeficiency virus (HIV) infection. Anti-
CC angiogenic therapy has little toxicity, does not require the therapeutic
CC agent to enter tumour cells or cross the blood-brain barrier, controls
CC tumour growth independently of growth of tumour cell heterogeneity, and
CC does not induce drug resistance
XX
SQ Sequence 1152 AA;
Query Match 93.2%; Score 55; DB 3; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 QGVQLQNVRFVF 12
Db 191 QGVQLQNVRFVF 201

Search completed: June 5, 2006, 22:24:58
Job time : 91.1379 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-25
Perfect score: 59
Sequence: 1 A QGVQLQNVRFVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80:*
1: Pirl1:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	93.2	229	2 S57957	thrombospondin 1 -
2	55	93.2	1170	1 TSHUP1	thrombospondin 1 p
3	55	93.2	1170	2 A40558	thrombospondin 1 p
4	37	62.7	151	2 C57253	tRNA-pseudouridine
5	37	62.7	1172	1 TSHUP2	thrombospondin 2 p
6	37	62.7	1172	2 A42587	thrombospondin 2 p
7	36	61.0	186	2 B75421	probable pilin, ty
8	35	59.3	40	2 T08870	alternative respir
9	35	59.3	89	2 B39529	cadherin-associate
10	35	59.3	102	2 D81910	hypothetical prote
11	35	59.3	326	2 T08849	alternative respir
12	35	59.3	747	2 AE2929	two component resp
13	35	59.3	783	2 A38353	probable transcrip
14	35	59.3	927	2 T43110	lactacin 481/lacto
15	35	59.3	1178	1 A39804	thrombospondin pre
16	34	57.6	102	2 G81105	hypothetical prote
17	34	57.6	106	2 C90261	hypothetical prote
18	34	57.6	265	2 T40878	probable FAD synth
19	34	57.6	308	2 JC5468	leukocidin chain 1
20	34	57.6	311	2 C89968	leukotoxin Luke [i
21	34	57.6	333	2 JC2432	L-lactate dehydrog
22	34	57.6	337	2 E37882	hypothetical prote
23	34	57.6	417	2 H83708	hypothetical prote
24	34	57.6	595	2 A43534	Lupus autoantigen
25	34	57.6	597	2 B82881	hypothetical prote
26	34	57.6	993	2 T17230	hypothetical prote
27	34	57.6	1308	2 T05178	hypothetical prote
28	33	55.9	54	2 S35637	leukocidin chain F
29	33	55.9	175	2 S75258	hypothetical prote

30 33 55.9 212 2 D81929 Probable imidazole
31 33 55.9 247 1 A64590 Probable 3-oxoacyl
32 33 55.9 247 2 B71923 3-oxoacyl-l-acyl-ca
33 33 55.9 286 2 C49238 gamma-hemolysin co
34 33 55.9 298 2 C87403 Fdhd protein limpo
35 33 55.9 310 2 S68225 synergohymenotropi
36 33 55.9 312 2 T00160 leukocidin chain S
37 33 55.9 312 2 S32211 leukocidin chain S
38 33 55.9 315 2 A49234 leukocidin R S com
39 33 55.9 315 2 JN0626 leukocidin chain S
40 33 55.9 315 2 E30043 gamma-hemolysin co
41 33 55.9 315 2 PC4078 hlgC-like protein
42 33 55.9 326 2 S54267 repA protein - Bac
43 33 55.9 331 2 S54263 rep A protein - Ba
44 33 55.9 350 1 RGBOT1 GTP-binding regula
45 33 55.9 350 1 RGHUT1 GTP-binding regula

ALIGNMENTS

RESULT 1

S57957
thrombospondin 1 - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S57957
R:Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.
submitted to the EMBL Data Library, July 1995
A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi
A:Reference number: S57955
A:Accession: S57957
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <LA>
A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:9899228; PTD
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; voi

Query Match 93.2%; Score 55; DB 2; Length 229;

Best Local Similarity 100.0%; Pred. No. 0.0017;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12

Db 191 QGVQLQNVRFVF 201

RESULT 2

TSHUP1
thrombospondin 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: A26155; A34274; A30140; A25812; A05172; A42927
R:Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c
A:Reference number: A26155; MUID:87057617; PMID:2430973
A:Accession: A26155
A:Molecule type: mRNA
A:Residues: 1-1170 <LA>
A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:937137; PIDN:CH
A:Note: parts of this sequence, including the amino end of the mature protein, were deter
R:Lahterty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA se
A:Reference number: A34274; MUID:89291870; PMID:2544587
A:Accession: A34274
A:Molecule type: DNA
A:Residues: 1-166 <LA>
A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835
R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I
J. Cell Biol. 108, 729-736, 1989
A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

A;Reference number: A30140; MUID:89139590; PMID:2918029
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-522,'A','524-1170 <HEN>
A;Cross-references: UNIPARC:UPI0000039AB1; EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Kobayashi, S.; Eden-McCuchan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis o
A;Reference number: A25812; MUID:87157592; PMID:3030396
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-397 <KOB>
A;Cross-references: UNIPARC:UPI000016B0CA; GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:
R;Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276; PMID:3461443
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-374,'RC' <DIX>
A;Cross-references: UNIPARC:UPI000016B140; GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A;Reference number: A42927; MUID:92348511; PMID:1379247
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Cross-references: UNIPARC:UPI00001742C0
A;Note: Cys-992 is shown to have a free sulphydryl
C;Genetics:
A;Gene: GDB:THBS1; TSP1; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <WVC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 93.2%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
| | | | | | | | | | | | | |
Db 209 QGVQLQNVRFVF 219

RESULT 3
A40558
Thrombospondin 1 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
R;Accession: A40558; A37905; B42587; S68787
R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991
A;Title: Characterization of the murine thrombospondin gene.
A;Reference number: A40558; MUID:92128941; PMID:1774063
A;Accession: A40558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1170 <LAW>
A;Cross-references: UNIPROT:P35441; UNIPARC:UPI0000028012; GB:M62449; GB:M62450; GB:M62451;
1; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469;
R;Bornstein, P.; Alfai, D.; Devarayalu, S.; Framson, P.; Li, P.
J. Biol. Chem. 265, 16691-16698, 1990
A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of
A;Reference number: A37905; MUID:90375546; PMID:2398070
A;Accession: A37905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <BOR>
A;Cross-references: UNIPARC:UPI000016D076; GB:J05605; GB:J05606; NID:g201991; PIDN:AAA404;
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: B42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1152,'P',1154-1170 <LAH>
A;Cross-references: UNIPARC:UPI0000177A96; GB:M87276
A;Note: sequence extracted from NCBI backbone (NCBI:P:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 ar
A;Reference number: S68787; MUID:96234006; PMID:8654563
A;Accession: S68787
A;Molecule type: protein
A;Residues: 13-26,'X',28-37 <CHE>
A;Cross-references: UNIPARC:UPI0000177A97
C;Complex: homotrimer, disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor
C;Keywords: calcium binding; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <WVC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
| | | | | | | | | | | | | |
Db 209 QGVQLQNVRFVF 219

RESULT 4
C57253
tRNA-pseudouridine synthase (EC 5.4.99.-) - Acinetobacter calcoaceticus (fragment)
N;Alternate names: hypothetical protein lipB 5'-region
C;Species: Acinetobacter calcoaceticus
C;Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 05-Oct-2004
C;Accession: C57253
R;Kok, R.G.; van Thor, J.J.; Nugteren-Roodzant, I.M.; Vosman, B.; Hellingwerf, K.J.
J. Bacteriol. 177, 3295-3307, 1995
A;Title: Characterization of lipase-deficient mutants of Acinetobacter calcoaceticus BD41
A;Reference number: A57253; MUID:95286514; PMID:7768830
A;Accession: C57253
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-151 <KOK>
A;Cross-references: UNIPARC:UPI000017879A; GB:X80800

C;Superfamily: tRNA pseudouridine synthase B
C;Keywords: intramolecular transferase; isomerase; tRNA modification

Query Match 62.7%; Score 37; DB 2; Length 151;
Best Local Similarity 58.3%; Pred. No. 5.2;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQGVQLQNVRFVF 12
:|||||:|:|:
Db 27 SNGVLQKVRWLF 38

RESULT 5
TSHUP2
thrombospondin 2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: A47379; A42173
R;LaBelli, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
A;Reference number: A47379; MUID:94010892; PMID:8406456
A;Accession: A47379
A;Molecule type: mRNA
A;Residues: 1-1172 <LAB>
A;Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046680; GB:L12350; NID:g307505; PIDN:
R;LaBelli, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression c
A;Reference number: A42173; MUID:92217961; PMID:1559694
A;Accession: A42173
A;Molecule type: mRNA
A;Residues: 560-1172 <LA2>
A;Cross-references: UNIPARC:UPI00001742C1; GB:M81339
A;Experimental source: fibroblast
A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBI:P:95096)
C;Genetics:
A;Gene: GDB:THBS2; TSP2
A;Cross-references: GDB:128789; OMIM:188061
A;Map position: 6q27-6q27
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1172/Product: thrombospondin 2 #status predicted <MAT>
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: thrombospondin type 1 repeat homology <EGF1>
F;652-691/Domain: EGF homology <EGF>
F;928-930/Region: cell attachment (R-G-D) motif
F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predi
F;167-226/Disulfide bonds: #status predicted
F;266,270/Disulfide bonds: interchain #status predicted
F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 62.7%; Score 37; DB 1; Length 1172;
Best Local Similarity 63.6%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
:|||||:|:
Db 203 RGLLQNVHLVF 213

RESULT 6
A42587
thrombospondin 2 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A42587; A39851
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell i
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: A42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1172 <LAH>
A;Cross-references: UNIPROT:Q03350; UNIPARC:UPI0000029847; GB:L07803; GB:M87275; NID:g34
A;Note: sequence extracted from NCBI backbone (NCBI:P:81502)
R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A;Title: A second, expressed thrombospondin gene (thbs2) exists in the mouse genome.
A;Reference number: A39851; MUID:91302287; PMID:1712771
A;Accession: A39851
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-873 <BOR>
A;Cross-references: UNIPARC:UPI000016D077; GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:5
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor
C;Keywords: calcium binding; glycoprotein
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>

Query Match 62.7%; Score 37; DB 2; Length 1172;
Best Local Similarity 63.6%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
:|||||:|:
Db 203 RGLLQNVHLVF 213

RESULT 7
B75421
probable pilin, type IV - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: B75421
F;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-186 <WHI>
A;Cross-references: UNIPROT:Q9RUZ7; UNIPARC:UPI00000D3E03; GB:AE001971; GB:AE000513; NID:
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1233
A;Map position: 1

Query Match 61.0%; Score 36; DB 2; Length 186;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNV 9
:|||||:|:
Db 60 QGVLENVR 67

RESULT 8
T08870
alternative respiratory pathway oxidase (EC 1.-.-.-) 3 - soybean (fragment)
C;Species: Glycine max (soybean)

C;Accession: A98353
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A98353
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-783 <KUR>
A;Cross-references: UNIPROT:Q8UB11; UNIPARC:UPI00000D2059; GB:AE007870; PIDN:AAK90347.1;
C;Genetics:
A;Gene: AGR_L3540
A;Map position: linear chromosome

Query Match 59.3%; Score 35; DB 2; Length 783;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GVLQNVRFV 11
|:|:|:|:|
Db 78 GILESVRFV 86

RESULT 14

T43110
lactacin 481/lactococin biosynthesis protein LCNDR2 - Lactococcus lactis plasmid pMRC01
C;Species: Lactococcus lactis
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T43110
R;Dougherty, B.A.; Hill, C.; Weidman, J.F.; Richardson, D.R.; Venter, J.C.; Ross, R.P. Mol. Microbiol. 29, 1029-1038, 1998
A;Title: Sequence and analysis of the 60 kb conjugative, bacteriocin-producing plasmid pMRC01
A;Reference number: Z22314
A;Accession: T43110
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-927 <DOU>
A;Cross-references: UNIPROT:O87240; UNIPARC:UPI00000B9A11; EMBL:AE001272; PIDN:AAC56013.
A;Experimental source: strain DPC3147
C;Genetics:
A;Genome: plasmid pMRC01
A;Note: ORF00039

Query Match 59.3%; Score 35; DB 2; Length 927;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AOGVLQNVRFV 12
|:|:|:|:|
Db 821 ASGVLQTLFFVY 832

RESULT 15

A39804
thrombospondin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A39804
R;Lawler, J.; Duquette, M.; Ferro, P. J. Biol. Chem. 266, 8039-8043, 1991
A;Title: Cloning and sequencing of chicken thrombospondin.
A;Reference number: A39804; MUID:91217026; PMID:2022631
A;Accession: A39804
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <LAW>
A;Cross-references: UNIPROT:P35440; UNIPARC:UPI000013776D; GB:M60853; NID:g212763; PIDN:
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vC
P;325-383/Domain: von Willebrand factor type C repeat homology <VWC>
P;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
P;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
P;499-555/Domain: thrombospondin type 1 repeat homology <THR3>

F;658-697/Domain: EGF homology <EGF>

Query Match 59.3%; Score 35; DB 1; Length 1178;
Best Local Similarity 45.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 OGVLQNVRFVF 12
|:|:|:|:|
Db 209 RGLLQNIHLIF 219

Search completed: June 5, 2006, 22:44:59
Job time : 14.9655 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 108.931 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-25
Perfect score: 59
Sequence: 1 AQGVQLQNVRFVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	93.2	229	2	Q28194 BOVIN
2	55	93.2	249	2	Q5U903_PIG
3	55	93.2	496	2	Q5Y84 XENLA
4	55	93.2	1170	1	TSP1 BOVIN
5	55	93.2	1170	1	P07996 homo sapien
6	55	93.2	1170	1	P35441 mus musculus
7	55	93.2	1170	2	Q3TRQ MOUSE
8	55	93.2	1170	2	Q71SA3 RAT
9	55	93.2	1171	2	Q80YQ1 MOUSE
10	55	93.2	1171	2	Q8CGB2 MOUSE
11	55	93.2	1173	1	TSP1 XENLA
12	55	93.2	1225	2	Q59E39 HUMAN
13	50	84.7	1090	2	Q5SPG5 BRARE
14	50	84.7	1171	2	Q4RLK5 TETNG
15	50	84.7	1193	2	Q4S758 TETNG
16	45	76.3	1034	2	Q4RQ74 TETNG
17	42	71.2	704	2	Q4P665 USTWA
18	42	71.2	1549	2	Q3CJ99 THEET
19	41	69.5	713	2	Q3FLU8 BURK
20	41	69.5	713	2	Q44XL2 BURK
21	41	69.5	713	2	Q4LLM8 BURK
22	40	67.8	378	2	Q563V1 XENLA
23	40	67.8	380	2	Q563S6 XENOPUS
24	40	67.8	380	2	Q563S7 GPIPI
25	40	67.8	380	2	Q563S8 GPIPI
26	40	67.8	380	2	Q563S9 GPIPI
27	40	67.8	380	2	Q563T0 GPIPI
28	40	67.8	380	2	Q563T1 GPIPI
29	40	67.8	380	2	Q563T2 GPIPI
30	40	67.8	380	2	Q563T3 GPIPI
31	40	67.8	380	2	Q563T4 XENOPUS

32	40	67.8	380	2	Q563T5 GPIPI	Q563T5 xenopus ami
33	40	67.8	380	2	Q563T6 GPIPI	Q563T6 xenopus pyg
34	40	67.8	380	2	Q563U0 XENCL	Q563U0 xenopus cili
35	40	67.8	380	2	Q563U1 GPIPI	Q563U1 xenopus bou
36	40	67.8	380	2	Q563U3 GPIPI	Q563U3 xenopus fra
37	40	67.8	380	2	Q563U4 GPIPI	Q563U4 xenopus fra
38	40	67.8	380	2	Q563U5 GPIPI	Q563U5 xenopus bou
39	40	67.8	380	2	Q563U7 GPIPI	Q563U7 xenopus ves
40	40	67.8	380	2	Q563U9 GPIPI	Q563U9 xenopus wit
41	40	67.8	380	2	Q563V0 GPIPI	Q563V0 xenopus wit
42	40	67.8	380	2	Q563V2 GPIPI	Q563V2 xenopus gil
43	40	67.8	443	2	Q2Y2B6 XNEOB	Q2Y2B6 chiralus
44	40	67.8	443	2	Q6J921 XNEOB	Q6J921 chiralus
45	40	67.8	523	2	Q70BX1 DROVI	Q70BX1 drosophila

ALIGNMENTS

RESULT 1
Q28194 BOVIN
ID Q28194 BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96331130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCPI9>3.3.CO;2-0;
RA Lafuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
J. Cell. Physiol. 167:164-172(1996).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; X89511; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:cell adhesion; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_g_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON TER 1
FT NON TER 229
FT NON TER 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9EBC4E6B669C CRC64;
Query Match 93.2%; Score 55; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 QGVQLQNVRFVF 12
Db 191 QGVQLQNVRFVF 201
RESULT 2
Q5U903_PIG
ID Q5U903_PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.

```

DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang K., Maucou G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY773342; AAV38110.1; -; mRNA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PSS0092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
FT NON_TER 249
FT NON_TER 249
SQ SEQUENCE 249 AA; 27560 MW; 465D664BE0329C32 CRC64;

Query Match 93.2%; Score 55; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
Db 1 QGVQLQNVRFVF 11

RESULT 3
Q7SY84 XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

```

```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC054970; AAH54970.1; -; mRNA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR GO; GO:0007155; P:Cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PSS0092; TSP1; 2.
DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE; PS50184; VWFC_2; 1.
SQ SEQUENCE 496 AA; 54643 MW; E4FD2F07CB7EF51B CRC64;

Query Match 93.2%; Score 55; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQNVRFVF 12
Db 215 QGVQLQNVRFVF 225

RESULT 4
TSPL_BOVIN
ID TSPL_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
RX MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
RT odontoblasts and predentin.";
RL Biochim. Biophys. Acta 1382:17-22(1998).

```

FT	DOMAIN	723	758	TSP type-3 1.				
FT	DOMAIN	759	781	TSP type-3 2.				
FT	DOMAIN	782	817	TSP type-3 3.				
FT	DOMAIN	818	840	TSP type-3 4.				
FT	DOMAIN	841	878	TSP type-3 5.				
FT	DOMAIN	879	914	TSP type-3 6.				
FT	DOMAIN	915	950	TSP type-3 7.				
FT	DOMAIN	951	1170	TSP C-terminal.				
FT	REGION	19	232	Heparin-binding (Potential).				
FT	MOTIF	926	928	Cell attachment site (Potential).				
FT	CARBOHYD	248	248	N-linked (GlcNAc. . .) (Potential).				
FT	CARBOHYD	360	360	N-linked (GlcNAc. . .) (Potential).				
FT	CARBOHYD	708	708	N-linked (GlcNAc. . .) (Potential).				
FT	CARBOHYD	1067	1067	N-linked (GlcNAc. . .) (Potential).				
FT	CARBOHYD	1085	1085	N-linked (GlcNAc. . .) (Potential).				
FT	DISULFID	270	270	Interchain (Probable).				
FT	DISULFID	274	274	Interchain (Probable).				
FT	DISULFID	391	423	By similarity.				
FT	DISULFID	395	428	By similarity.				
FT	DISULFID	406	413	By similarity.				
FT	DISULFID	447	484	By similarity.				
FT	DISULFID	451	489	By similarity.				
FT	DISULFID	462	474	By similarity.				
FT	DISULFID	504	541	By similarity.				
FT	DISULFID	508	546	By similarity.				
FT	DISULFID	519	531	By similarity.				
FT	DISULFID	551	562	By similarity.				
FT	DISULFID	556	572	By similarity.				
FT	DISULFID	575	586	By similarity.				
FT	DISULFID	592	608	By similarity.				
FT	DISULFID	599	617	By similarity.				
FT	DISULFID	620	644	By similarity.				
FT	DISULFID	650	663	By similarity.				
FT	DISULFID	657	676	By similarity.				
FT	DISULFID	678	689	By similarity.				
FT	DISULFID	705	713	By similarity.				
FT	DISULFID	718	738	By similarity.				
FT	DISULFID	754	774	By similarity.				
FT	DISULFID	777	797	By similarity.				
FT	DISULFID	813	833	By similarity.				
FT	DISULFID	836	856	By similarity.				
FT	DISULFID	874	894	By similarity.				
FT	DISULFID	910	930	By similarity.				
FT	DISULFID	946	1167	By similarity.				
FT	CONFLICT	805	805	S -> G (in Ref. 2).				
SEQ	SEQUENCE	1170 AA;	129534 MW;	ODD6ADF3E5FA031A CRC64;				
Query Match 93.2%; Score 55; DB 1; Length 1170;								
Best Local Similarity 100.0%; Pred.No. 0.12;								
Matches 11; Conservative 0; Mismatches 0; Indels 0;								
Qy	2 QGVQLNVRPFV 12							
Db	209 QGVQLNVRPFV 219							
RESULT 5								
TSP1_HUMAN								
ID	TSP1_HUMAN	STANDARD;	PRT;	1170 AA.				
AC	P07996; Q15667;							
CD	01-AUG-1988, integrated into UniProtKB/Swiss-Prot.							
DT	01-AUG-1988, sequence version 1.							
DT	07-MAR-2006, entry version 78.							
DE	Thrombospondin-1 precursor.							
GN	Name=THBS1; Synonyms=TFP, TSP1;							
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homi							
OC	Homo.							
OX	NCBI_TaxID=9606;							
RN	[1]							
RP	NUCLEOTIDE SEQUENCE.							
RC	TISSUE=Endothelial cell;							

XX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologies with several different
RT proteins.";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RX Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region.";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RN NUCLEOTIDE SEQUENCE OF 1-397.
RP MEDLINE=87157592; PubMed=3030396;
RX Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
RA "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins.";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RN NUCLEOTIDE SEQUENCE OF 1-374.
RP MEDLINE=86287276; PubMed=3461443;
RX Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RA "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RN NUCLEOTIDE SEQUENCE OF 1-166.
RP MEDLINE=89291870; PubMed=2544587;
RX Laherty C.D., German T.M., Dixit V.M.;
RA "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RN NUCLEOTIDE SEQUENCE OF 1028-1170.
RP la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RA "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RN CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;
RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT module.";
RL J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RN THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RP MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RX Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RA "Biophysical characterization, including disulfide bond assignments,
RT of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
RN [9]
RN CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RP PubMed=16335952; DOI=10.1021/pr0502065;
RX Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hydrazide chemistry, and mass spectrometry.";
RL J. Proteome Res. 4:2070-2080(2005).
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- SIMILARITY: Belongs to the thrombospondin family.

CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; M25631; AAA36741.1; -; mRNA.
CC EMBL; X04665; CAA28370.1; -; mRNA.
CC EMBL; X14787; CAA32889.1; -; mRNA.
CC EMBL; M14326; AAA61237.1; ALT SEQ; mRNA.
CC EMBL; J04835; AAA61178.1; -; Genomic_DNA.
CC EMBL; M99425; AAB59366.1; -; mRNA.
CC PIR; A26155; TSHUPL.
CC PDB; 1LSL; X-ray; A=434-546.
CC PDB; 1UX6; X-ray; A=834-1170.
CC PDB; 1Z78; X-ray; A=19-233.
CC PDB; 1ZA4; X-ray; A=19-257.
CC PDB; 2ERF; X-ray; A=25-233.
CC SMR; P07996; 549-1169.
CC GlycoSuiteDB; P07996; -.
CC OGP; P07996; -.
CC Ensembl; ENSG00000137801; Homo sapiens.
CC HGNC; HGNC:11785; THBS1.
CC MIM; 188060; gene.
CC Reactome; P07996; -.
CC GO; GO:0005576; C:extracellular region; NAS.
CC GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
CC GO; GO:0004871; F:signal transducer activity; TAS.
CC GO; GO:0007275; P:development; TAS.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSP_C.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR001007; VWFC_C.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC Pfam; PF00093; VWC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS01208; VWFC_1; 1.
CC PROSITE; PS0184; VWFC_2; 1.
CC 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
CC Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO_0000035842.
FT TSP N-terminal.
FT DOMAIN 24 221 VWFC.
FT DOMAIN 316 373 TSP type-1 1.
FT DOMAIN 379 429 TSP type-1 2.
FT DOMAIN 435 490 TSP type-1 3.
FT DOMAIN 492 547 EGF-like 1.
FT DOMAIN 549 587 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 588 645


```

FT DOMAIN 646 690 EGF-like 3.
FT 723 758 TSP type-3 1.
FT 759 781 TSP type-3 2.
FT 782 817 TSP type-3 3.
FT 818 840 TSP type-3 4.
FT 841 878 TSP type-3 5.
FT 879 914 TSP type-3 6.
FT 915 950 TSP type-3 7.
FT 951 1170 TSP C-terminal.
FT REGION 19 232 Heparin-binding (Potential).
FT MOTIF 926 928 Cell attachment site (Potential).
FT CARBOHYD 248 248 N-linked (GlcNAc...).
FT CARBOHYD 360 360 N-linked (GlcNAc...). (Potential).
FT CARBOHYD 385 385 C-linked (Man).
FT CARBOHYD 394 394 /FTid=CAR 000205.
FT CARBOHYD 438 438 /FTid=CAR 000206.
FT CARBOHYD 441 441 C-linked (Man).
FT CARBOHYD 450 450 /FTid=CAR 000207.
FT CARBOHYD 450 450 C-linked (Man).
FT CARBOHYD 498 498 /FTid=CAR 000208.
FT CARBOHYD 507 507 /FTid=CAR 000209.
FT CARBOHYD 507 507 C-linked (Man).
FT CARBOHYD 507 507 /FTid=CAR 000210.
FT CARBOHYD 507 507 O-linked (Fuc...).
FT CARBOHYD 708 708 /FTid=CAR 000211.
FT CARBOHYD 1067 1067 N-linked (GlcNAc...). (Potential).
FT DISULFID 270 270 N-linked (GlcNAc...).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 391 423 Interchain (Probable).
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546

Query Match 93.2%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVFV 12
Db 209 QGVQLQVRVFV 219

RESULT 6
TSPI MOUSE
ID TSPI MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 57.
DE Thrombospondin-1 precursor.
GN Name=Thbs1; Synonyms=Tspl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
[2]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=92147683; PubMed=1371115;

```

```

RA Laberty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression
RT during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-490.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfi D., Devarayalu S., Framson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of
RT the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
RN [4]
RP PROTEIN SEQUENCE OF 19-37.
RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
RT "Expression and initial characterization of recombinant mouse
RT thrombospondin 1 and thrombospondin 3.";
RL FEBS Lett. 387:36-41(1996).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 WFEC domain.
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; M62470; AAA50611.1; -; Genomic DNA.
DR EMBL; M62450; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62451; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62452; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62453; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62454; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62455; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62456; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62457; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62458; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62459; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62460; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62461; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62462; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62463; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62464; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62465; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62466; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62467; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62468; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62469; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M87276; AAA5063.1; -; mRNA.
DR EMBL; J05606; AAA40431.1; -; Genomic DNA.
DR EMBL; J05605; AAA40431.1; JOINED; Genomic DNA.
DR PIR; A40558; A40558.
DR HSP; P07996; 1LSL.
DR SMR; P35441; 549-1169.
DR Ensembl; ENSMUSG00000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.

```

DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; TSP_3.
 DR InterPro; IPR008859; TSP_C.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 12.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; WVC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; WVC_1; 1.
 DR PROSITE; PS00184; WVC_2; 1.
 DR Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
 KW Glycoprotein; Heparin-binding; Repeat; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1170
 FT DOMAIN 24 221
 FT DOMAIN 316 373
 FT DOMAIN 379 429
 FT DOMAIN 435 490
 FT DOMAIN 492 547
 FT DOMAIN 549 587
 FT DOMAIN 588 645
 FT DOMAIN 646 690
 FT DOMAIN 723 758
 FT DOMAIN 759 781
 FT DOMAIN 782 817
 FT DOMAIN 818 840
 FT DOMAIN 841 878
 FT DOMAIN 879 914
 FT DOMAIN 915 950
 FT DOMAIN 951 1170
 FT REGION 19 232
 FT MOTIF 926 928
 FT CARBOHYD 248 248
 FT CARBOHYD 360 360
 FT CARBOHYD 708 708
 FT CARBOHYD 1067 1067
 FT DISULFID 270 270
 FT DISULFID 274 274
 FT DISULFID 391 423
 FT DISULFID 395 428
 FT DISULFID 406 413
 FT DISULFID 447 484
 FT DISULFID 451 489
 FT DISULFID 462 474
 FT DISULFID 504 541
 FT DISULFID 508 546
 FT DISULFID 519 531
 FT DISULFID 551 562
 FT DISULFID 556 572
 FT DISULFID 575 586
 FT DISULFID 592 608
 FT DISULFID 599 617
 FT DISULFID 620 644
 FT DISULFID 650 663
 FT DISULFID 657 676
 FT DISULFID 678 689
 FT DISULFID 705 713
 FT DISULFID 718 738
 FT DISULFID 754 774
 FT DISULFID 777 797
 FT DISULFID 813 833

FT DISULFID 836 856 By similarity.
 FT DISULFID 874 894 By similarity.
 FT DISULFID 910 930 By similarity.
 FT DISULFID 946 1167 By similarity.
 FT CONFLICT 1025 1025 F -> L (in Ref. 2).
 SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;
 Query Match 93.2%; Score 55; DB 1; Length 1170;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 QGVQLQNVRFVF 12
 Db 209 QGVQLQNVRFVF 219
 RESULT 7
 Q3TR40 MOUSE
 ID Q3TR40 MOUSE PRELIMINARY; PRT; 1170 AA.
 AC Q3TR40;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 5.
 DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
 DE clone:AS30055N06 product:thrombospondin 1, full insert sequence.
 GN Name=Thbs1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 [1]
 RP NUCLEOTIDE SEQUENCE
 RC STRAIN=CS7BL/6J; TISSUE=Aorta and vein;
 RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RL Methods Enzymol. 303:19-44(1999).
 [2]
 RP NUCLEOTIDE SEQUENCE
 RC STRAIN=CS7BL/6J; TISSUE=Aorta and vein;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
 Ambesi-Impombato A., Apweiler R., Auraliysa R.N., Bailey T.L.,
 Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
 Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
 Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,
 Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sinclair B.,
 Shibata Y., Shimada H., Shimada K., Sultana R., Takenaka Y., Taki K.,
 Sperling S., Stupka E., Sugiuira K., Sultana R., Tegner J., Teichmann S.A.,
 Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 Yamanishi H., Zabarovskiy E., Zhu S., Zimmer A., Hide W., Bult C.,
 Grimmond S.M., Teasdale R.D., Liu E.T., Brusica V., Quackenbush J.,
 Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,

RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nagami T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Balzarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher T.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Garibolchi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nomata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kogawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikolaev I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balzarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).

RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Onoda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
EMBL; AK163092; BA037190.1; -; mRNA.
MG1; MGI:98737; Thbs1.
GO; GO:0005615; C:extracellular space; IDA.
GO; GO:0005615; C:extracellular space; RCA.
GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
InterPro; IPR006210; EGF.
InterPro; IPR000742; EGF_3.
InterPro; IPR001881; EGF_Ca_bd.
InterPro; IPR006209; EGF_like.
InterPro; IPR013032; EGF_like_reg.
InterPro; IPR003129; Laminin_G_TSP_N.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP_1.
InterPro; IPR003367; tsp_3.
InterPro; IPR008859; TSP_C.
InterPro; IPR001007; VWF_C.
Pfam; PF00008; EGF; 2.
Pfam; PF00090; TSP_1; 3.
Pfam; PF02412; TSP_3; 12.
Pfam; PF05735; TSP_C; 1.
Pfam; PF00093; VMC; 1.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00181; EGF; 3.
SMART; SM00209; TSP1; 3.
SMART; SM00210; TSPN; 1.
SMART; SM00214; VMC; 1.
PROSITE; PS01186; EGF_2; UNKNOWN_1.
PROSITE; PS0026; EGF_3; 2.
PROSITE; PS0092; TSP1; 3.
PROSITE; PS01208; VWF_C; 1.
PROSITE; PS0184; VWF_C; 2; 1.
SQ SEQUENCE 1170 AA; 129619 MW; 55BF04A20D91B194 CRC64;
Query Match 93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

```
Qy 2 QGVQLQVRFFV 12
Db 209 QGVQLQVRFFV 219

RESULT 8
Q71SA3 RAT PRELIMINARY; PRT; 1170 AA.
AC Q71SA3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Thrombospondin 1.
GN Name=Tsp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RA Iwabu A., Hirohata S., Kusachi S., Nakamura K., Murakami T.,
RA Ninomiya Y., Tsuji T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AF309630; AAQ14549.1; -; mRNA.
DR SRR; Q71SA3; 834-1169.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR06210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_Tsp_N.
DR InterPro; IPR008084; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSEN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE; PS0184; VWFC_2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRFFV 12
Db 209 QGVQLQVRFFV 219
```

```
RESULT 9
Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Haie S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; BC050917; AAH50917.1; -; mRNA.
DR HSP; P07996; 1LSL.
DR SRR; Q80YQ1; 835-1170.
DR Ensembl; ENSMUSG00000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR06210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_Tsp_N.
DR InterPro; IPR008084; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
```


RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walshstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zvolanik M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arawaka T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
RT [7]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arawaka T., Hara A., Fukuhashi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuoka H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriber L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690 (2001).
RT [8]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RT Genome Res. 10:1617-1630 (2000).
RT [9]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multipipette sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RT [10]
RN NUCLEOTIDE SEQUENCE.
RN TISSUE=Mammary gland;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watah

```

CC EMBL; BC042422; AAH42422.1; -; mRNA.
DR EMBL; AK145202; BAE26293.1; -; mRNA.
DR HSSP; P07996; 1LSL.

Query Match          93.2%   Score 55; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps
Matches 11; Conservative 0; Mismatches 0;

Qy      2 QGVQLQVRVFV 12
        |||||
Db      209 QGVQLQVRVFV 219

RESULT 11
TS1P1_XENLA
ID _TS1P1_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 54.
DE Thrombospondin-1 precursor.
DE Name=thbs1; Synonyms=tspl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RT "Cloning, characterization, and expression of thrombospondin-1 in
RL Xenopus laevis embryos.";
CC Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-cell and
CC cell-to-matrix interactions. Can bind to fibronogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.

Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
HSSP; P07996; 1LSL.
SMR; P35448; 552-1172.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.

```

DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF 1; FALSE NEG.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00026; EGF 3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWC; 1.
DR PROSITE; PS01208; VWC; 1.
DR PROSITE; PS01208; VWC; 1.
DR PROSITE; PS01208; VWC; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1173
FT DOMAIN 23 224
FT DOMAIN 319 376
FT DOMAIN 382 432
FT DOMAIN 438 493
FT DOMAIN 495 550
FT DOMAIN 550 590
FT DOMAIN 591 648
FT DOMAIN 649 693
FT DOMAIN 726 761
FT DOMAIN 762 784
FT DOMAIN 785 820
FT DOMAIN 821 843
FT DOMAIN 844 881
FT DOMAIN 882 917
FT DOMAIN 918 953
FT DOMAIN 954 1173
FT REGION 23 335
FT MOTIF 929 931
FT CARBOHYD 155 155
FT CARBOHYD 158 158
FT CARBOHYD 250 250
FT CARBOHYD 363 363
FT CARBOHYD 705 705
FT CARBOHYD 711 711
FT CARBOHYD 1070 1070
FT CARBOHYD 394 426
FT DISULFID 398 431
FT DISULFID 409 416
FT DISULFID 450 487
FT DISULFID 454 492
FT DISULFID 465 477
FT DISULFID 507 544
FT DISULFID 511 549
FT DISULFID 522 534
FT DISULFID 554 565
FT DISULFID 559 575
FT DISULFID 578 589
FT DISULFID 595 611
FT DISULFID 602 620
FT DISULFID 623 647
FT DISULFID 653 666
FT DISULFID 660 679
FT DISULFID 681 692
FT DISULFID 708 716
FT DISULFID 721 741
FT DISULFID 757 777
FT DISULFID 780 800
FT DISULFID 816 836
FT DISULFID 839 859
FT DISULFID 877 897
FT DISULFID 913 933
FT DISULFID 949 1170
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 93.2%; Score 55; DB 1; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQVRVFV 12
|||||

Db 212 QGVQLQVRVFV 222
RESULT 12
Q59E99 HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RA "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC
DR EMBL; AB209912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939, 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00026; EGF 3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWC 1; 1.
DR PROSITE; PS01208; VWC 2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9880B16E57157B12 CRC64;

Query Match 93.2%; Score 55; DB 2; Length 1225;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQVRVFV 12
|||||

Db 264 QGVQNVRFVF 274

```
RESULT 13
QSSPG5_BRARE PRELIMINARY; PRT; 1090 AA.
AC QSSPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (by similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AL928866; CA120599.1; -; Genomic DNA.
DR SMR; QSSPG5; 751-804, 754-1089.
DR Ensembl; ENSDARG0000010785; Danio rerio.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR002048; EF hand_Ca_bd.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF 3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR006884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF HAND 1; UNKNOWN_1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWF_C; 1; 1.
DR PROSITE; PS0184; VWF_C; 2; 1.
DR Cell adhesion; EGF-like domain.
FT NON TER
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;
```

Query Match 84.7%; Score 50; DB 2; Length 1090;
Best Local Similarity 100.0%; Pred.No.1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GVLQNVRFVF 12
| | | | | | | |
Db 187 GVLQNVRFVF 196

```
RESULT 14
Q4RLR5_TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Chromosome 10 SCAF15019, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00032374001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=95883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castella S.,
RA Anthouard V., Jubin C., Castelli V., Kankina M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Parra G., Lardier S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAE01015019; CAG10667.1; -; Genomic DNA.
DR SMR; Q4RLR5; 834-887, 837-1171.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF 3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR03129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR01007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP 1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
```



```

DR SMART; SM00209; TSPI; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSPI; 3.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS50184; VWF_2; 1.
KW Cell adhesion.
FT NON TER 1171 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 84.7%; Score 50; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GVLQNVRFVF 12
DB 209 GVLQNVRFVF 218

RESULT 15
Q4S758 TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, entry version 8.
DE Chromosome 14 SCAP14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG0022976001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Kankina M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; CAAG01014723; CAG03524.1; -; Genomic_DNA.
DR SMR; Q4S758; 811-1148.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.

```

```

DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1RPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSPI; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSPI; 3.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS50184; VWF_2; 1.
KW Cell adhesion.
FT NON TER 1 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 1193;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GVLQNVRFVF 12
DB 189 GVLQNVRFVF 198

Search completed: June 5, 2006, 22:42:39
Job time : 109.931 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-25

Perfect score: 59

Sequence: 1 AQGVQLQVRVFV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /EMC Cellerai_SIDS3/ptodata/2/iaa/5 COMB.pdp.*
 - 2: /EMC Cellerai_SIDS3/ptodata/2/iaa/6 COMB.pdp.*
 - 3: /EMC Cellerai_SIDS3/ptodata/2/iaa/7 COMB.pdp.*
 - 4: /EMC Cellerai_SIDS3/ptodata/2/iaa/H COMB.pdp.*
 - 5: /EMC Cellerai_SIDS3/ptodata/2/iaa/PCTUS COMB.pdp.*
 - 6: /EMC Cellerai_SIDS3/ptodata/2/iaa/RE COMB.pdp.*
 - 7: /EMC Cellerai_SIDS3/ptodata/2/iaa/backfilea1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	93.2	825	2	US-09-949-002-482
2	55	93.2	831	2	US-09-939-853A-97
3	55	93.2	831	2	US-09-939-853A-98
4	55	93.2	1170	1	US-08-313-288B-20
5	55	93.2	1170	2	US-09-657-472-2
6	55	93.2	1170	2	US-09-949-002-350
7	39	66.1	731	2	US-09-270-767-42057
8	37	62.7	446	2	US-09-199-637A-267
9	37	62.7	1045	2	US-09-949-016-11112
10	37	62.7	1172	1	US-08-313-288B-19
11	37	62.7	1172	2	US-09-949-016-6333
12	35	59.3	175	2	US-09-252-991A-18826
13	34	57.6	89	2	US-09-270-767-46041
14	34	57.6	231	2	US-09-605-703B-44
15	34	57.6	277	2	US-09-605-703B-42
16	34	57.6	390	2	US-08-650-766-7
17	34	57.6	390	2	US-08-922-635-6
18	34	57.6	390	2	US-09-328-352-4891
19	34	57.6	390	2	US-09-389-487-7
20	34	57.6	390	2	US-09-414-643-6
21	34	57.6	559	2	US-09-364-206-47
22	34	57.6	651	2	US-08-650-766-6
23	34	57.6	651	2	US-08-922-635-5
24	34	57.6	651	2	US-09-389-487-6
25	34	57.6	651	2	US-09-414-643-5
26	34	57.6	1070	2	US-08-922-635-22

27	34	57.6	1070	2	US-09-414-643-22	Sequence 22, Appl
28	34	57.6	1504	2	US-09-364-206-2	Sequence 2, Appl
29	33	55.9	40	1	US-07-868-353A-3	Sequence 3, Appl
30	33	55.9	40	1	US-08-407-804-3	Sequence 3, Appl
31	33	55.9	40	2	US-09-124-807-3	Sequence 3, Appl
32	33	55.9	135	2	US-09-270-767-33170	Sequence 33170, A
33	33	55.9	135	2	US-09-270-767-48387	Sequence 48387, A
34	33	55.9	273	2	US-09-710-279-84	Sequence 84, Appl
35	33	55.9	273	2	US-09-134-001C-3732	Sequence 3732, Ap
36	33	55.9	350	1	US-07-868-353A-14	Sequence 14, Appl
37	33	55.9	350	1	US-08-407-804-23	Sequence 23, Appl
38	33	55.9	350	2	US-09-124-807-23	Sequence 23, Appl
39	33	55.9	354	1	US-07-868-353A-12	Sequence 12, Appl
40	33	55.9	354	1	US-07-868-353A-13	Sequence 13, Appl
41	33	55.9	354	1	US-07-868-353A-15	Sequence 15, Appl
42	33	55.9	354	1	US-08-407-804-21	Sequence 21, Appl
43	33	55.9	354	1	US-08-407-804-22	Sequence 22, Appl
44	33	55.9	354	1	US-08-407-804-24	Sequence 24, Appl
45	33	55.9	354	2	US-09-124-807-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human

Query Match

Best Local Similarity 93.2%; Score 55; DB 2; Length 825;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVFV 12

Db 309 QGVQLQVRVFV 319

RESULT 2

US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:

; APPLICANT: Buttgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          93.2%; Score 55; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
      |||||
Db      209 QGVQLQNVRFVF 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232zel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          93.2%; Score 55; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
      |||||
Db      209 QGVQLQNVRFVF 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          93.2%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
      |||||
Db      209 QGVQLQNVRFVF 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
      |||||
Db      209 QGVQLQNVRFVF 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
; US-09-949-002-350

Query Match          93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 QGVQLQNVRFV 12
Db  209 QGVQLQNVRFV 219

RESULT 7
US-09-270-767-42057
; Sequence 42057, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42057
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-42057

Query Match          66.1%; Score 39; DB 2; Length 731;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  2 QGVQLQNVRFV 11
Db  238 QGVQLQNVDFM 247

RESULT 8
US-09-199-637A-267
; Sequence 267, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
```

```
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 267
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-267

Query Match          62.7%; Score 37; DB 2; Length 446;
Best Local Similarity 72.7%; Pred. No. 44;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy  1 AQGVQLQNVRFV 11
Db  329 AQGALENVRKV 339

RESULT 9
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-11112

Query Match          62.7%; Score 37; DB 2; Length 1045;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy  2 QGVQLQNVRFV 12
Db  282 RGLQLQNVHLVF 292

RESULT 10
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match 62.7%; Score 37; DB 1; Length 1172;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQNVRVVF 12
Db 203 RGLQNHLVF 213

RESULT 11

US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6333

Query Match 62.7%; Score 37; DB 2; Length 1172;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQNVRVVF 12
Db 203 RGLQNHLVF 213

RESULT 12

US-09-252-991A-18826
; Sequence 18826, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18826
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18826

Query Match 59.3%; Score 35; DB 2; Length 175;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQNVRVVF 12
Db 41 QGVLDVQVFLF 51

RESULT 13

US-09-270-767-46041
; Sequence 46041, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46041
; LENGTH: 89
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-46041

Query Match 57.8%; Score 34; DB 2; Length 89;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LQNVRFVF 12
Db 78 LQNVRLVF 85

RESULT 14

US-09-605-703B-44
; Sequence 44, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habethauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 44
; LENGTH: 231
; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-44

Query Match 57.6%; Score 34; DB 2; Length 231;
Best Local Similarity 58.3%; Pred. No. 82;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AOGVLQNVRFVF 12
Db 143 AMGGLGSIRFVF 154

RESULT 15

US-09-605-703B-42
; Sequence 42, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605.703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 42
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-42

Query Match 57.6%; Score 34; DB 2; Length 277;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AOGVLQNVRFVF 12
Db 189 AMGGLGSIRFVF 200

Search completed: June 5, 2006, 22:48:53
Job time : 23.8966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-25
Perfect score: 59
Sequence: 1 AQGVQNVRPFV 12

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main.*
1: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	93.2	12	4	US-10-474-213-28
2	55	93.2	240	4	Sequence 28, Appl
3	55	93.2	240	5	Sequence 40, Appl
4	55	93.2	432	5	Sequence 1020, Ap
5	55	93.2	432	5	Sequence 1022, Ap
6	55	93.2	459	6	Sequence 462, App
7	55	93.2	466	3	Sequence 1047, Ap
8	55	93.2	555	6	Sequence 454, App
9	55	93.2	578	6	Sequence 456, App
10	55	93.2	685	6	Sequence 452, App
11	55	93.2	804	6	Sequence 453, App
12	55	93.2	828	6	Sequence 455, App
13	55	93.2	831	3	Sequence 97, Appl
14	55	93.2	831	3	Sequence 98, Appl
15	55	93.2	855	6	Sequence 461, App
16	55	93.2	1000	6	Sequence 457, App
17	55	93.2	1105	6	Sequence 459, App
18	55	93.2	1150	4	Sequence 1, Appli
19	55	93.2	1152	3	Sequence 1, Appli
20	55	93.2	1169	5	Sequence 7, Appli
21	55	93.2	1170	4	Sequence 12, Appl
22	55	93.2	1170	4	Sequence 2, Appli
23	55	93.2	1170	4	Sequence 114, App
24	55	93.2	1170	4	Sequence 2, Appli
25	55	93.2	1170	4	Sequence 1170, Ap
26	55	93.2	1170	4	Sequence 38, Appl
27	55	93.2	1170	4	Sequence 482, Appl

28	55	93.2	1170	4	US-10-419-462-38	Sequence 38, Appl
29	55	93.2	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	55	93.2	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	55	93.2	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	55	93.2	1170	5	US-10-782-968-38	Sequence 38, Appl
33	55	93.2	1170	5	US-10-849-989-44	Sequence 44, Appl
34	55	93.2	1170	5	US-10-631-467-548	Sequence 548, App
35	55	93.2	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	55	93.2	1170	5	US-10-831-997-2	Sequence 2, Appli
37	55	93.2	1170	5	US-10-995-561-594	Sequence 594, App
38	55	93.2	1170	5	US-10-995-561-595	Sequence 595, App
39	55	93.2	1170	5	US-10-995-561-596	Sequence 596, App
40	55	93.2	1170	6	US-11-037-713-51	Sequence 51, Appl
41	55	93.2	1170	6	US-11-046-644-28	Sequence 28, Appl
42	55	93.2	1170	6	US-11-046-456-28	Sequence 28, Appl
43	39	66.1	15	4	US-10-285-394-153	Sequence 153, App
44	39	66.1	721	6	US-11-097-143-13287	Sequence 13287, A
45	39	66.1	3500	4	US-10-153-219-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Krutzsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/10535
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 93.2%; Score 55; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQNVRPFV 12
Db 2 QGVQNVRPFV 12

RESULT 2
US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40

```
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match          93.2%; Score 55; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QGVQLQNVRFVF 12
Db      191 QGVQLQNVRFVF 201

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match          93.2%; Score 55; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QGVQLQNVRFVF 12
Db      191 QGVQLQNVRFVF 201

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match          93.2%; Score 55; DB 5; Length 432;
```

```
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

RESULT 5
US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match          93.2%; Score 55; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

RESULT 6
US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match          93.2%; Score 55; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

RESULT 7
US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-03-925-301-1047

Query Match 93.2%; Score 55; DB 3; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVF 12
| | | | | | | | | |
Db 262 QGVQLQVRVF 272

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 93.2%; Score 55; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVF 12
| | | | | | | | | |
Db 209 QGVQLQVRVF 219

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 93.2%; Score 55; DB 6; Length 578;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVF 12
| | | | | | | | | |
Db 209 QGVQLQVRVF 219

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 93.2%; Score 55; DB 6; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVF 12
| | | | | | | | | |
Db 209 QGVQLQVRVF 219

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 93.2%; Score 55; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVQLQVRVF 12
| | | | | | | | | |
Db 209 QGVQLQVRVF 219

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match          93.2%; Score 55; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          93.2%; Score 55; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-09-939-853A-98

Query Match          93.2%; Score 55; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match          93.2%; Score 55; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFVF 12
Db      209 QGVQLQNVRFVF 219

Search completed: June 6, 2006, 00:00:10
Job time : 79.6207 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-25

Perfect score: 59

Sequence: 1 AQLVQNVRFV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pap:*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pap:*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pap:*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pap:*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	57.6	182	6	US-10-953-349-15318
2	34	57.6	608	7	US-11-203-828-5
3	34	57.6	1504	6	US-10-505-928-662
4	33	55.9	298	6	US-10-953-349-9347
5	33	55.9	331	6	US-10-953-349-34674
6	33	55.9	358	6	US-10-953-349-34673
7	33	55.9	382	6	US-10-953-349-34672
8	32	54.2	196	6	US-10-953-349-5024
9	32	54.2	207	6	US-10-953-349-5023
10	32	54.2	273	6	US-10-953-349-5022
11	32	54.2	285	6	US-10-953-349-26582
12	32	54.2	297	7	US-11-293-697-4132
13	32	54.2	440	6	US-10-953-349-1348
14	31	52.5	855	7	US-11-247-437-2
15	31	52.5	1234	6	US-10-953-349-3997
16	31	52.5	1237	6	US-10-953-349-3996
17	31	52.5	1337	6	US-10-953-349-3995
18	30	50.8	74	6	US-10-953-349-15283
19	30	50.8	77	6	US-10-953-349-15282
20	30	50.8	197	6	US-10-953-349-30699
21	30	50.8	197	6	US-10-953-349-31888
22	30	50.8	210	6	US-10-953-349-21828
23	30	50.8	223	6	US-10-953-349-21827
24	30	50.8	227	6	US-10-953-349-39138
25	30	50.8	265	6	US-10-953-349-39137

26	30	50.8	273	6	US-10-953-349-21826	Sequence 21826, A
27	30	50.8	276	6	US-10-953-349-39136	Sequence 39136, A
28	30	50.8	358	6	US-10-953-349-30698	Sequence 30698, A
29	30	50.8	358	6	US-10-953-349-31887	Sequence 31887, A
30	30	50.8	407	6	US-10-953-349-26200	Sequence 26200, A
31	30	50.8	459	6	US-10-953-349-7070	Sequence 7070, Ap
32	30	50.8	470	6	US-10-953-349-7069	Sequence 7069, Ap
33	30	50.8	1018	7	US-11-293-697-2998	Sequence 2998, Ap
34	30	50.8	1043	6	US-10-511-937-2452	Sequence 2452, Ap
35	29.5	50.0	574	7	US-11-121-154-49	Sequence 49, Appl
36	29	49.2	139	7	US-11-293-697-3323	Sequence 3323, Ap
37	29	49.2	262	6	US-10-953-349-2388	Sequence 2388, Ap
38	29	49.2	290	6	US-10-953-349-5462	Sequence 5462, Ap
39	29	49.2	296	6	US-10-953-349-2387	Sequence 2387, Ap
40	29	49.2	329	6	US-10-953-349-5461	Sequence 5461, Ap
41	29	49.2	364	6	US-10-953-349-11888	Sequence 11888, A
42	29	49.2	437	6	US-10-504-120-32	Sequence 32, Appl
43	29	49.2	525	6	US-10-953-349-31207	Sequence 31207, A
44	29	49.2	560	6	US-10-953-349-31206	Sequence 31206, A
45	29	49.2	656	6	US-10-953-349-1717	Sequence 1717, Ap

ALIGNMENTS

RESULT 1

US-10-953-349-15318
; Sequence 15318, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15318
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15318

Query Match 57.6%; Score 34; DB 6; Length 182;
Best Local Similarity 60.0%; Pred. No. 7.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GVLQNVRFV 12
Db 104 GLLMMNFVF 113

RESULT 2

US-11-203-828-5
; Sequence 5, Application US/11203828
; Publication No. US20060110390A1
; GENERAL INFORMATION:
; APPLICANT: SUCHAROV, CARMEN
; APPLICANT: LEINWAND, LESLIE
; TITLE OF INVENTION: INHIBITION OF KU AS A TREATMENT FOR CARDIOVASCULAR DISEASES
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: MYOG:58US
; CURRENT APPLICATION NUMBER: US/11/203.828
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/604,435
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus

```
US-11-203-828-5
Query Match          57.6%; Score 34; DB 7; Length 608;
Best Local Similarity 45.5%; Pred. No. 26;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVQLQVRFFV 12
: : : : :
Db 461 KAIVQVRFTY 471

RESULT 3
US-10-505-928-662
; Sequence 662, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 662
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-662

Query Match          57.6%; Score 34; DB 6; Length 1504;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQGVQLQVRFFV 12
: : : : :
Db 750 SQHILSLRFFV 761

RESULT 4
US-10-953-349-9347
; Sequence 9347, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9347

Query Match          55.9%; Score 33; DB 6; Length 298;
Best Local Similarity 40.0%; Pred. No. 19;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVQLQVRFFV 11
: : : : :
Db 87 EGIIQGVKFI 96

RESULT 5
US-10-953-349-34674
; Sequence 34674, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match          55.9%; Score 33; DB 6; Length 331;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVQLQVRFFV 11
: : : : :
Db 225 QGVLFNIQYV 234

RESULT 6
US-10-953-349-34673
; Sequence 34673, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

Query Match          55.9%; Score 33; DB 6; Length 358;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVQLQVRFFV 11
: : : : :
Db 252 QGVLFNIQYV 261

RESULT 7
US-10-953-349-34672
; Sequence 34672, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34672
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34672

Query Match          55.9%; Score 33; DB 6; Length 382;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2  QGVLRNRFV 11
      ||||| |::|
Db      276  QGVLFNIQV 285

RESULT 8
US-10-953-349-5024
; Sequence 5024, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5024
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5024

Query Match      54.2%; Score 32; DB 6; Length 196;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy      1  AQGV--LQNVRFV 12
      ||::| |::|
Db      34  ARGVYLSNIRVF 47

RESULT 9
US-10-953-349-5023
; Sequence 5023, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5023
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5023

Query Match      54.2%; Score 32; DB 6; Length 207;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy      1  AQGV--LQNVRFV 12
      ||::| |::|
Db      45  ARGVYLSNIRVF 58

RESULT 10
US-10-953-349-5022
; Sequence 5022, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
```

```
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5022
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5022

Query Match      54.2%; Score 32; DB 6; Length 273;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy      1  AQGV--LQNVRFV 12
      ||::| |::|
Db      111  ARGVYLSNIRVF 124

RESULT 11
US-10-953-349-26582
; Sequence 26582, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26582
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-26582

Query Match      54.2%; Score 32; DB 6; Length 285;
Best Local Similarity 58.3%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      1  AQGVLRNRFV 12
      ||||| |::|
Db      6  ADGERQNDFVF 17

RESULT 12
US-11-293-697-4132
; Sequence 4132, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; PRIOR FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4132
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4132

Query Match      54.2%; Score 32; DB 7; Length 297;
Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      3  GVLQNVRFV 12
      ||::| |::|
Db      234  GVFNRVRFV 243
```

```
RESULT 13
US-10-953-349-1348
; Sequence 1348, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579FUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1348
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1348

Query Match          54.2%; Score 32; DB 6; Length 440;
Best Local Similarity 60.0%; Pred.No. 45;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      3  GVQLQNVRFV 12
      |||:|:|:|
Db      182  GVLENSAIF 191

RESULT 14
US-11-247-437-2
; Sequence 2, Application US/11247437
; Publication No. US2006010753A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Roberts, Bruce
; TITLE OF INVENTION: METHODS TO DIAGNOSE AND TREAT LUNG CANCER
; FILE REFERENCE: 5257C
; CURRENT APPLICATION NUMBER: US/11/247,437
; CURRENT FILING DATE: 2005-10-11
; PRIOR APPLICATION NUMBER: PCT/US2004/011193
; PRIOR FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US 60/462,028
; PRIOR FILING DATE: 2003-04-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-247-437-2

Query Match          52.5%; Score 31; DB 7; Length 855;
Best Local Similarity 50.0%; Pred.No. 1.5e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      2  QGVQLQNVRFV 11
      :|:|:|:|:|
Db      445  RGVYENVKIV 454

RESULT 15
US-10-953-349-3997
; Sequence 3997, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579FUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
```

```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3997
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3997

Query Match          52.5%; Score 31; DB 6; Length 1234;
Best Local Similarity 55.6%; Pred.No. 2.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      4  VLQNVRFV 12
      :|:|:|:|
Db      1122  ILENVRLVY 1130

Search completed: June 6, 2006, 00:12:55
Job time : 3.82414 secs
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-26

Perfect score: 60

Sequence: 1 FAGVLQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_8.1

1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*
10: Geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	12	4	AAB35369 Alpha3bet
2	55	91.7	12	4	AAB35352 Alpha3bet
3	55	91.7	12	4	AAB35378 Alpha3bet
4	55	91.7	12	6	ABG72834 Thrombos
5	55	91.7	240	8	ADL70641 Human thr
6	55	91.7	432	8	ADQ39359 Human myo
7	55	91.7	432	8	ADQ39357 Human myo
8	55	91.7	459	4	AAU02916 Angiotens
9	55	91.7	466	3	ABA43602 Human can
10	55	91.7	546	4	AAU02915 Angiotens
11	55	91.7	548	7	ADN02474 TSF polyp
12	55	91.7	555	4	AAU02914 Angiotens
13	55	91.7	731	4	AAU02913 Angiotens
14	55	91.7	1152	3	ABA00042 Human thr
15	55	91.7	1152	5	AAU74771 Human thr
16	55	91.7	1152	5	ABB82285 Human thr
17	55	91.7	1170	4	AAB74450 Human var
18	55	91.7	1170	4	AAB90800 Human she
19	55	91.7	1170	5	AAE25030 Human thr
20	55	91.7	1170	5	AAU75315 Human thr
21	55	91.7	1170	6	ABP96780 Human COP
22	55	91.7	1170	6	ABU03474 Angiogene
23	55	91.7	1170	6	ABG74673 Human THB

24	55	91.7	1170	6	AAE36228 Human THB
25	55	91.7	1170	7	ABR62059 Human thr
26	55	91.7	1170	7	ADN39852 Cancer/an
27	55	91.7	1170	8	ADJ76124 Marker ge
28	55	91.7	1170	8	ADJ75296 Marker ge
29	55	91.7	1170	8	ADL70639 Human thr
30	55	91.7	1170	8	ADL35874 Human thr
31	55	91.7	1170	8	ADQ26070 Thrombos
32	55	91.7	1170	8	ADP54179 Human PKO
33	55	91.7	1170	8	ADQ39358 Human myo
34	55	91.7	1170	8	ADQ39356 Human myo
35	55	91.7	1170	8	ADQ39355 Human myo
36	55	91.7	1170	9	ADZ21688 Thrombos
37	55	91.7	1170	9	ABE87781 Human thr
38	55	91.7	1170	9	ABE46751 Human thr
39	52	86.7	12	4	AAAB35373 Alpha3bet
40	52	86.7	12	4	AAAB35381 Alpha3bet
41	51	85.0	12	4	AAAB35364 Alpha3bet
42	51	85.0	12	4	AAAB35374 Alpha3bet
43	50	83.3	11	4	AAAB35357 Alpha3bet
44	50	83.3	12	4	AAAB35371 Alpha3bet
45	50	83.3	12	4	AAAB35368 Alpha3bet

ALIGNMENTS

RESULT 1

AAB35369
ID AAB35369 standard; peptide; 12 AA.

AC AAB35369;

DT 08-MAY-2001 (first entry)

DE Alpha3beta1 integrin binding peptide #34.

KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;

KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;

KW macular degeneration; psoriasis; cell adhesion; cell motility.

OS Synthetic.

PN WO200105812-A2.

PD 25-JAN-2001.

PF 12-JUL-2000; 2000WO-US018986.

PR 15-JUL-1999; 99US-0144549P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Roberts DD, Kruttsch HC;

DR WPI; 2001-182656/18.

PT New peptides that bind to or are recognized by alpha3-beta1 integrins, useful for inhibiting cell adhesion to extracellular matrix, cell motility and proliferation and for treating rheumatoid arthritis and cancer.

PT Claim 4; Page 34; 84pp; English.

CC The present invention provides a number of peptides which bind to alpha3beta1 integrins. They are useful in the modulation of cell adhesion and motility, and in the treatment of cancer, diabetic retinopathy, rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis and restenosis. The present sequence is an example of one of the peptides of the invention

SQ Sequence 12 AA;

Query Match 100.0%; Score 60; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 1 FAGVLQNVRFVF 12

RESULT 2
AAB35352
ID AAB35352 standard; peptide; 12 AA.
XX
AC AAB35352;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3betal integrin binding peptide #17.
XX
KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
FN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;
Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0007;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 1 FQGVQLQNVRFVF 12

RESULT 3
AAB35378
ID AAB35378 standard; peptide; 12 AA.
XX
AC AAB35378;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3betal integrin binding peptide #43.

XX Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
FN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Example 2; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;
Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0007;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 1 FQGVQLQNVRFVF 12

RESULT 4
ABG72834
ID ABG72834 standard; peptide; 12 AA.
XX
AC ABG72834;
XX
DT 24-FEB-2003 (first entry)
XX
DE Thrombospondin-1 sequence containing synthetic peptide.
XX
KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
OS Synthetic.
XX
FN WO200281630-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010535.
XX
PR 06-APR-2001; 2001US-0281994P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;

```

XX WPI; 2003-103329/09.
XX
PT A new diagnosis for cancer other than prostate cancer in a mammal useful
XX to detect cancer including lung cancer, particularly small cell lung
XX cancer and melanoma comprises detecting semenogelin in a sample.
XX
PS Example 1; Page 14; 32pp; English.
XX
CC The invention relates to diagnosing cancer other than prostate cancer in
CC a male mammal, comprising assaying a test sample for increased level of
CC semenogelin, or cancer in a female by assaying for the presence of
CC semenogelin. Administering a semenogelin protein or polypeptide fragment
CC or a semenogelin-specific antibody or active fragment, or a recombinant
CC vector expressing the protein or antibody, is useful for inducing an
CC immune response to a cancer in a mammal, where the cancer is not prostate
CC cancer and semenogelin is a marker. The invention is used to diagnose
CC cancer, particularly of epithelial origin such as lung cancer, papillary
CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
CC (SCLC), or a melanoma. The present sequence represents the amino acid
CC sequence of the thrombospondin-1 sequence containing synthetic peptide
CC which binds to alpha-3-beta-1 integrin
XX
SQ Sequence 12 AA;
Query Match 91.7%; Score 55; DB 6; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0007;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FAGVLQNVRFVF 12
DB 1 FQGVQLQNVRFVF 12
RESULT 5
ADL70641
ID ADL70641 standard; protein; 240 AA.
AC ADL70641;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human thrombospondin-1 N-terminal domain.
XX
XX Human; thrombospondin-1; epitope; cancer; diagnosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 23..32
XX /note= "Heparin binding region"
XX Region 77..82
XX /note= "Heparin binding region"
XX Region 151..164
XX /note= "Fibrinogen binding region"
XX
XX WO2004018995-A2.
XX
XX 04-MAR-2004.
XX
XX 20-AUG-2003; 2003WO-US026023.
XX
XX 23-AUG-2002; 2002US-0405494P.
XX
XX 21-APR-2003; 2003US-00419462.
XX
XX (WILL/) WILLIAMS K J.
XX
XX Williams KJ;
XX
XX WPI; 2004-226901/21.
XX
XX New purified thrombospondin fragment extracted from a body fluid, useful
XX for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
XX

```

```

PT or leukemia or as calibrators, indicators, immunogens and analytes.
XX
PS Disclosure; SEQ ID NO 40; 76pp; English.
XX
XX The present sequence is that of the N-terminal domain of human
XX thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
XX (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
XX clinical assays for cancer and for generation of antibodies and other
XX binding agents. A method that distinguishes TSP from a TSP fragment or
XX portion involves: (1) using an epitope shared by TSP and the TSP fragment
XX or portion as a target for a binding molecule, e.g. an antibody, to
XX obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
XX epitope present in TSP but not in the fragment or portion to obtain a
XX quantitation of TSP only; and (3) using the difference between (1) and
XX (2) as a quantitation of the amount of TSP fragment or portion. Suitable
XX epitopes are provided ADL70602-ADL70638. Detection or quantification of
XX the TSP fragment or portion is performed in order to detect the presence,
XX or monitor the course, of a disease or condition selected from cancer,
XX renal failure, renal disease, atopic dermatitis, vasculitis, acute
XX vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
XX infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
XX nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
XX syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
XX rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
XX thrombocytopenic purpura, purpura, an inflammatory condition, a condition
XX associated with clotting, a condition associated with platelet
XX activation, a condition associated with intravascular platelet
XX activation, a condition associated with consumption of platelets, heparin
XX -induced thrombocytopenia, disseminated intravascular coagulation,
XX intravascular coagulation, extravascular coagulation, a condition
XX associated with endothelial activation, a condition associated with
XX production and/or release of thrombospondin and/or a thrombospondin
XX fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
XX reaction, an aspartame reaction, atopic dermatitis, eczema,
XX hypersensitivity, scleroderma, conditions associated with plugging of
XX vessels, a condition associated with a cryofibrinogen, a condition
XX associated with a cryoglobulin, and a condition associated with an anti-
XX cardiolin antibody. The cancer is selected from adenoma,
XX adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
XX cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
XX a cancer with vascular invasion, internal cancer, skin cancer, cancer of
XX the respiratory system, circulatory system, musculoskeletal system,
XX muscle, bone, a joint, tendon or ligament, digestive system, liver or
XX biliary system, pancreas, head, neck, endocrine system, reproductive
XX system (male or female), genitourinary system, kidney, urinary tract,
XX sensory system, nervous system, lymphoid organ, blood, a gland, mammary
XX gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
XX tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
XX well-differentiated cancer or a moderately differentiated cancer.
XX
SQ Sequence 240 AA;
Query Match 91.7%; Score 55; DB 8; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FAGVLQNVRFVF 12
DB 190 FQGVQLQNVRFVF 201
RESULT 6
ADQ39359
ID ADQ39359 standard; protein; 432 AA.
XX
XX ADQ39359;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX cardiac; gene therapy; human.
XX

```

XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38531.
XX PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1022; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;
Query Match 91.7%; Score 55; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FAGVLQNVRFVF 12
| | | | | | | | | | | | | |
Db 208 FQGVQLQNVRFVF 219
RESULT 7
ADQ39357
ID ADQ39357 standard; protein; 432 AA.
XX AC ADQ39357;
XX DT 18-NOV-2004 (first entry)

XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1020.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human.
XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38529.
XX PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1020; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;
Query Match 91.7%; Score 55; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FAGVLQNVRFVF 12
| | | | | | | | | | | | | |
Db 208 FQGVQLQNVRFVF 219
RESULT 8
AAU02916

ID	AAU02916 standard; protein; 459 AA.	ID	AAB43602 standard; protein; 466 AA.
XX	AAU02916;	XX	AAB43602;
AC		AC	
XX		XX	
XX		XX	
DT	12-SEP-2001 (first entry)	DT	08-FEB-2001 (first entry)
XX		XX	
DE	Angiotensin converting enzyme (ACEV) splice variant protein #16.	DE	Human cancer associated protein sequence SEQ ID NO:1047.
XX		XX	
KW	Angiotensin converting enzyme splice variant; ACEV; interleukin 6;	KW	Human; cancer associated gene; cancer antigen; detection; cancer;
KW	granulocyte colony stimulating factor receptor; glucagon; hypertrophy;	KW	diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
KW	platelet-derived endothelial cell growth factor; cardiovascular disease;	KW	antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW	cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;	KW	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW	vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;	KW	dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
KW	myocardial infarction; coronary arterial thrombosis; renal disease;	KW	vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;
KW	diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;	KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW	multiple sclerosis; immune complex nephritis; deep vein thrombosis;	KW	allergic reaction; graft versus host disease; organ rejection;
KW	nonrheumatic pulmonary granulomatous disease; endothelial abnormality;	KW	haemostatic; thrombolytic; cardiovascular disorder; infection;
KW	vascular disorder; asbestosis.	KW	neurological disease; drug screening.
XX		XX	
OS	Homo sapiens.	XX	Homo sapiens.
XX		OS	
XX		XX	
PN	WO200136632-A2.	XX	WO200055350-A1.
XX		PN	
XX		XX	
PD	25-MAY-2001.	XX	21-SEP-2000.
XX		PD	
XX		XX	
PF	17-NOV-2000; 2000WO-IL000766.	XX	08-MAR-2000; 2000WO-US005882.
XX		PF	
XX		XX	
PR	17-NOV-1999; 99IL-00132978.	XX	12-MAR-1999; 99US-0124270P.
XX		PR	
PR	10-DEC-1999; 99IL-00133455.	XX	(HUMA-) HUMAN GENOME SCI INC.
XX		XX	
XX	(COMP-) COMPUGEN LTD.	PA	
PA		XX	
XX		XX	
XX		PI	Rosen CA, Ruben SM;
PI	Levine Z, David A, Azar I, Khosravi R, Bernstein J;	XX	
XX		XX	
XX		DR	WPI; 2000-587533/55.
DR	WPI; 2001-336004/35.	XX	N-PSDB; AAC77811.
XX		DR	
XX		XX	
PT	Novel alternative splicing variants e.g. variant of angiotensin	PT	Novel isolated nucleic acids comprising sequences encoding peptides
PT	converting enzyme (ACEV), useful in identifying candidate compounds	PT	useful for treating or diagnosing e.g. cancer.
PT	capable of binding to the variant and to detect anti-variant antibodies.	XX	
XX		PS	Claim 11; Page 1636-1638; 2352pp; English.
PS	Claim 4; Fig 16; 519pp; English.	XX	
XX		CC	AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC	The sequence represents an angiotensin converting enzyme splice variant	CC	AAB43398 to AAB44239. The proteins can have activities based on the
CC	(ACEV) polypeptide. The polypeptides of the invention include variants of	CC	tissues and cells the genes are expressed in. Example of activities
CC	granulocyte colony stimulating factor receptor, glucagon, interleukin 6,	CC	include: cytostatic; proliferative; vulnerable; immunomodulator;
CC	platelet-derived endothelial cell growth factor, cyclin-dependent kinase	CC	antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC	inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal	CC	antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC	polypeptide receptor 2. The polypeptides and their associated nucleic	CC	dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC	acids are useful for identification of variant sequences and detection of	CC	neotropic; vasotropic; antipsoriatic and angiogenic. The
CC	candidate compounds capable of binding to the molecules. The sequences of	CC	polynucleotides and polypeptides can be used for preventing, treating or
CC	the invention can be used in the treatment and diagnosis of various	CC	ameliorating medical conditions and diagnosing pathological conditions.
CC	disorders including cardiovascular diseases such as arteriosclerosis,	CC	Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC	myocardial infarction and coronary arterial thrombosis, renal diseases	CC	the present invention may be used to treat immune disorders by activating
CC	such as diabetic nephropathy, muscular diseases such as hypertrophy,	CC	or inhibiting the proliferation, differentiation or mobilisation of
CC	immune disorders such as immune complex nephritis, multiple sclerosis,	CC	immune cells, to treat disorders of haematopoietic cells, autoimmune
CC	cancer, sarcoidosis, nonrheumatic pulmonary granulomatous diseases such	CC	disorders, allergic reactions, graft versus host disease and organ
CC	as asbestosis and vascular pathologies involving an endothelial	CC	rejection, modulate haemostatic or thrombolytic activity, modulate
CC	abnormality such as deep vein thrombosis	CC	inflammation, cancers, cardiovascular disorders, neurological disease and
XX		CC	bacterial or viral infections. The peptides, nucleotides, antibodies,
XX		CC	agonists and antagonists may be also be used in drug screens. AAC78449 to
XX		CC	AAC78457 and AAB44240 represent sequences used in the exemplification of
XX		CC	the present invention
SQ	Sequence 459 AA;	XX	Sequence 466 AA;
Query Match	91.7%; Score 55; DB 4; Length 459;	Query Match	91.7%; Score 55; DB 3; Length 466;
Best Local Similarity	91.7%; Pred. No. 0.04;	Best Local Similarity	91.7%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1 FAGVLQNVRFVF 12	Qy	1 FAGVLQNVRFVF 12
Db	208 FQGVQLQNVRFVF 219	Db	261 FQGVQLQNVRFVF 272
RESULT 9			
AAB43602			

RESULT 10
AAU02915
ID AAU02915 standard; protein; 546 AA.
XX
AC AAU02915;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX
OS Homo sapiens.
XX
FN WO200136632-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-IL000766.
XX
PR 17-NOV-1999; 99IL-00132978.
XX
PR 10-DEC-1999; 99IL-00133455.
XX
PA (COMP-) COMPUGEN LTD.
XX
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX
XX WPI; 2001-336004/35.
XX
DR N-PSDB; AAS06015.
XX
XX Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
XX capable of binding to the variant and to detect anti-variant antibodies.
XX
XX Claim 4; Fig 15; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
XX polypeptide receptor 2. The polypeptides and their associated nucleic
XX acids are useful for identification of variant sequences and detection of
XX candidate compounds capable of binding the molecules. The sequences of
XX the invention can be used in the treatment and diagnosis of various
XX disorders including cardiovascular diseases such as arteriosclerosis,
XX myocardial infarction and coronary arterial thrombosis, renal diseases
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX immune disorders such as immune complex nephritis, multiple sclerosis,
XX cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
XX as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
XX
SQ Sequence 546 AA;
XX
Query Match 91.7%; Score 55; DB 4; Length 546;
Best Local Similarity 91.7%; Pred. No. 0.048;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 FAGVLQNRVVF 12
| | | | | | | | | |
Db 208 FQGLQNRVVF 219
XX
RESULT 12
AAU02914
ID AAU02914 standard; protein; 555 AA.
XX
AC AAU02914;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW

RESULT 11
ADN02474
ID ADN02474 standard; protein; 548 AA.
XX
AC ADN02474;
XX
DT 17-JUN-2004 (first entry)
XX
DE TSF polypeptide.
XX
KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX
OS Homo sapiens.
XX
PN CNI401387-A.
XX
PD 12-MAR-2003.
XX
PF 21-AUG-2002; 2002CN-00129408.
XX
PR 21-AUG-2002; 2002CN-00129408.
XX
PA (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
XX
PI Han Z, Liu P;
XX
DR WPI; 2003-469302/45.
XX
DR N-PSDB; ADN02475.
XX
PT Tumor suppressing polypeptide TSF and gene therapy vector composition.
XX
PS Claim 2; SEQ ID NO 1; 13pp; Chinese.
XX
XX The present invention relates to a novel recombinant adenovirus vector
XX mediated anti-neoplastic composition is prepared through cloning the cDNA
XX sequence from the human peripheral blood cell by specific primer and
XX reverse transcription-polymerase chain reaction (RT-PCR) method for
XX coding TSF polypeptide, construction in human embryonic kidney cell 293
XX by AdEasy system, and packaging and expressing the recombinant adenovirus
XX vector of TSF. It can suppress the growth and transfer of cancer. The
XX present sequence represents the TSF polypeptide.
XX
SQ Sequence 548 AA;
XX
Query Match 91.7%; Score 55; DB 7; Length 548;
Best Local Similarity 91.7%; Pred. No. 0.048;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 FAGVLQNRVVF 12
| | | | | | | | | |
Db 208 FQGLQNRVVF 219
XX
RESULT 12
AAU02914
ID AAU02914 standard; protein; 555 AA.
XX
AC AAU02914;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW

KW vascular disorder; asbestosis.
XX

OS Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

PR 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

DR N-PSDB; AAS06014.

XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.

PS Claim 4; Fig 14; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis

XX SQ Sequence 555 AA;

Query Match 91.7%; Score 55; DB 4; Length 555;

Best Local Similarity 91.7%; Pred. No. 0.049;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12

Db 208 FQGVQLQNVRFVF 219

RESULT 13

AAU02913

ID AAU02913 standard; protein; 731 AA.

XX AAU02913;

XX 12-SEP-2001 (first entry)

XX Angiotensin converting enzyme (ACEV) splice variant protein #13.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

KW vascular disorder; asbestosis.

XX Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

PR 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

DR N-PSDB; AAS06013.

XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.

PS Claim 4; Fig 13; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis

XX SQ Sequence 731 AA;

Query Match 91.7%; Score 55; DB 4; Length 731;

Best Local Similarity 91.7%; Pred. No. 0.067;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12

Db 208 FQGVQLQNVRFVF 219

RESULT 14

AAB00042

ID AAB00042 standard; protein; 1152 AA.

XX AAB00042;

XX 08-NOV-2000 (first entry)

XX Human thrombospondon-1 (TSP-1).

XX TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
KW diabetic retinopathy; corneal graft rejection; glaucoma.

XX Homo sapiens.

XX Key Location/Qualifiers

XX 361..416

XX Region /label= Type 1 repeat region

FT Region 417..473
FT /label= Type 1 repeat region
FT 474..530
FT /label= Type 1 repeat region
XX
XX WO200044908-A2.
XX
XX 03-AUG-2000.
XX
XX 01-FEB-2000; 2000WO-US002482.
XX
XX 01-FEB-1999; 99US-0118053P.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX Lawler JW;
XX
XX WPI; 2000-514823/46.
XX
XX Nucleic acids encoding chimeric proteins such as cartilage oligomeric
PT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
PT inhibiting angiogenesis and treating diseases such as cancer.
XX
XX
XX Disclosure; Fig 1; 40pp; English.
XX
XX New nucleic acids are described which encode a protein comprising the
CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
CC the TGF (transforming growth factor)-beta activation region of human TSP-
CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
CC tumours in mice models. Thus the nucleic acids and proteins may be useful
CC for treating angiogenesis related diseases such as cancer (by reducing
CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
CC for treating human immunodeficiency virus (HIV) infection. Anti-
CC angiogenic therapy has little toxicity, does not require the therapeutic
CC agent to enter tumour cells or cross the blood-brain barrier, controls
CC tumour growth independently of growth of tumour cell heterogeneity, and
CC does not induce drug resistance
XX
XX Sequence 1152 AA;
SQ
Query Match 91.7%; Score 55; DB 3; Length 1152;
Best Local Similarity 91.7%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | | | | | |
Db 190 FQGVLQNVRFVF 201
RESULT 15
AAU74771
ID AAU74771 standard; protein; 1152 AA.
XX
AC AAU74771;
XX
XX 09-APR-2002 (first entry)
XX
XX Human thrombospondin-1 (TSP-1).
XX
XX Thrombospondin-1; TSP-1; cytostatic; angiogenesis; vasotropic; vulneryary;
KW neovascularisation; cell proliferation inhibitor; cancer; solid tumour;
KW haemangioma; acoustic neuromas; neurofibroma; trachoma;
KW pyogenic granulomas; rheumatoid arthritis; ocular angiogenic disease;
KW retinopathy; psoriasis; macular degeneration; corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; rebeosis; angiofibroma;
KW Osler-Webber syndrome; myocardial angiogenesis; haemophilic joints;
KW plaque neovascularisation; telanglectasia; wound granulation; apoptosis.
XX
XX Homo sapiens.
Qs

XX Key Location/Qualifiers
FH Region 263..360
FT /label= Procollagen_homology_region
FT 303..309
FT /label= Procollagen_homology_domain
FT /note= "Required in inhibition of angiogenesis"
FT 361..530
FT /label= Type 1 repeat domain
FT /note= "This region contains 3 type 1 repeats, from
FT residues 361-416, residues 417-473 and residues 474-530"
FT 364..370
FT /label= Heparin_binding_domain
FT 413..415
FT /label= RPK motif
FT /note= "Necessary and sufficient for activation of
FT transforming growth factor beta (TGF beta)."
FT 418..423
FT /label= TGF-beta and fibronectin_binding_domain
FT /note= "Transforming growth factor"
FT 420..426
FT /label= Heparin_binding_domain
FT 429..434
FT /label= Cell_binding_domain
FT 477..483
FT /label= Heparin_binding_domain
FT 481..499
FT /label= Anti-angiogenesis_domain
FT 486..491
FT /label= Cell_binding_domain
FT 531..673
FT /label= Type 2 repeat domain
FT /note= "This region contains 3 type 2 repeats, from
FT residues 531-571, residues 572-629 and residues 630-673"
FT 570..601
FT /label= Calcium_binding_domain
FT 698..925
FT /label= Type 3 repeat domain
FT /note= "This region contains 7 type 3 repeats, from
FT residues 698-733, residues 734-756, residues 757-792,
FT residues 793-815, residues 816-853, residues 854-889 and
FT residues 890-925"
XX
XX WO200191781-A2.
XX
XX 06-DEC-2001.
XX
XX 25-MAY-2001; 2001WO-US017250.
XX
XX 26-MAY-2000; 2000US-0207994P.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX Lawler JW;
XX
XX WPI; 2002-106273/14.
XX
XX Composition useful for treatment of cancer comprises cDNA encoding amino
PT acids of human thrombospondin-1 or its conservative variant and a
PT carrier.
XX
XX Disclosure; Fig 7; 54pp; English.
XX
XX The invention describes a composition comprising cDNA encoding fragments
CC of human thrombospondin-1 (TSP -1), a type 1 repeat polypeptide and its
CC potent inhibitor of tumour growth and angiogenesis. The composition is
CC useful for killing cancerous cells (preferably tumour); for reducing
CC volume or inhibiting growth of a tumour (inhibiting neovascularisation in
CC the tumour); for decreasing proliferation of tumour cells; in the
CC treatment of diseases and conditions associated with angiogenic activity
CC or misregulated growth and angiogenesis-mediated diseases such as cancer,
CC solid tumour, tumour metastasis, benign tumour, (e.g. haemangioma,
CC acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas),

CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases (e.g.
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasias,
CC rebeosis), Osler-Webber syndrome, myocardial angiogenesis,
CC telangiectasia, plaque neovascularisation, haemophilic joints,
CC angiofibroma or wound granulation. The composition induces apoptosis and
CC inhibits neovascularisation in the tumour cells. This amino acid sequence
CC represents human thrombospondin-1 (TSP-1), on which the recombinant
XX proteins of the invention are based

SQ Sequence 1152 AA;

Query Match 91.7%; Score 55; DB 5; Length 1152;

Best Local Similarity 91.7%; Pred. No. 0.11;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12

| | | | | | | | | |

Db 190 FQGVLQNVRFVF 201

Search completed: June 5, 2006, 22:25:00
Job time : 91.1379 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-26

Perfect score: 60

Sequence: 1 FAGVLQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1.*
- 2: Pir2.*
- 3: Pir3.*
- 4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	91.7	229	2 S57957	thrombospondin 1 -
2	55	91.7	1170	1 TSHUP1	thrombospondin 1 p
3	55	91.7	1170	2 A40558	thrombospondin 1 p
4	41	68.3	1172	1 TSHUP2	thrombospondin 2 p
5	41	68.3	1172	2 A42587	thrombospondin 2 p
6	39	65.0	747	2 A82929	two component resp
7	39	65.0	783	2 A98353	probable transcrip
8	39	65.0	1178	1 A39804	thrombospondin pre
9	37	61.7	145	2 T20985	hypothetical prote
10	37	61.7	162	2 B88349	protein F15D4.3 [i
11	37	61.7	780	2 T50315	hypothetical prote
12	36	60.0	151	2 C57253	tRNA-pseudouridine
13	36	60.0	300	2 D81399	malate dehydrogena
14	36	60.0	355	2 T50479	G protein alpha ch
15	36	60.0	440	2 F96556	IAA-Ala hydrolase
16	36	60.0	453	2 AE0774	probable protease
17	36	60.0	453	2 H64974	hypothetical prote
18	36	60.0	453	2 A98990	hypothetical prote
19	36	60.0	453	2 D85835	hypothetical prote
20	36	60.0	464	2 AG0347	probable proteinase
21	36	60.0	467	2 D84938	H ⁺ -transporting tw
22	36	60.0	492	2 S32491	testosterone 7alph
23	36	60.0	587	2 T16867	probable cytochrom
24	36	60.0	4572	2 S57908	hypothetical 527K
25	35	58.3	167	2 A11167	hypothetical prote
26	35	58.3	292	2 S60950	probable membrane
27	35	58.3	334	2 A83225	binding protein co
28	35	58.3	338	2 S56333	carbon-phosphorus
29	35	58.3	338	2 D86105	hypothetical prote

30	35	58.3	338	2	G91264	hypothetical prote
31	35	58.3	417	2	H83708	hypothetical prote
32	35	58.3	435	2	H96556	auxin conjugate hy
33	35	58.3	465	2	S76464	hypothetical prote
34	35	58.3	471	2	S45068	53K glycoprotein -
35	35	58.3	3587	2	I40486	surfactin syntheta
36	34	56.7	133	2	B89969	enterotoxin Yentl
37	34	56.7	167	2	AD1526	hypothetical prote
38	34	56.7	199	1	G64070	imidazoleglycerol-
39	34	56.7	308	2	JCS468	leukocidin chain 1
40	34	56.7	311	2	C89968	leukotoxin LtkE [i
41	34	56.7	337	2	E97882	hypothetical prote
42	34	56.7	459	2	D86669	amino acid permeas
43	34	56.7	460	2	G64066	collagenase prtC h
44	34	56.7	466	2	F82288	probable proteinase
45	34	56.7	541	2	S43061	t-complex-type mol

ALIGNMENTS

RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C/Accession: S57957

R/Lafeuilleade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

A:Description: Opposite regulation of thrombospondin-1 and Cisp/thrombospondin-2 expressi

A:Reference number: S57955

A/Accession: S57957

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-229 <LAF>

A/Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:g899228; PIDN:

C/Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

Query Match 91.7%; Score 55; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0037;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12

Db 190 FQGLQNVRFVF 201

RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)

C>Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C/Accession: A26155; A34274; A30140; A25812; A05172; A42927

R/Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A/Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c

A/Reference number: A26155; MUID:87057617; PMID:2430973

A/Accession: A26155

A/Molecule type: mRNA

A/Residues: 1-1170 <LAW>

A/Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:g37137; PIDN:CF

A/Note: parts of this sequence, including the amino end of the mature protein, were deter

R/Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A/Title: Characterization of the promoter region of the human thrombospondin gene. DNA se

A/Reference number: A34274; MUID:89291870; PMID:2544587

A/Accession: A34274

A/Molecule type: DNA

A/Residues: 1-166 <LAH>

A/Cross-references: UNIPARC:UPI00001742BF; GB:J04835

R/Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I

J. Cell Biol. 108, 729-736, 1989

A/Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

A;Reference number: A30140; MUID:89139590; PMID:2918029
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-522','A','524-1170 <HEN>
A;Cross-references: UNIPARC:UPI0000038AB1; EMBL:X14787; NID:G37464; PIDN:CAA32889.1; PIDN:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P. Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of complementary DNA
A;Reference number: A25812; MUID:87157592; PMID:3030396
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-397 <KOB>
A;Cross-references: UNIPARC:UPI00001680CA; GB:M25631; NID:G538353; PIDN:AAA36741.1; PIDN:Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A. Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276; PMID:3461443
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83,'A','85-374','RC' <DIX>
A;Cross-references: UNIPARC:UPI000016B140; GB:M14326; NID:G340005; PIDN:AAA61237.1; PIDN:Sun, X.; Skorstengaard, K.; Mosher, D.F. J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A;Reference number: A42927; MUID:92348511; PMID:1379247
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Cross-references: UNIPARC:UPI00001742C0
A;Note: Cys-992 is shown to have a free sulphydryl
C;Genetics:
A;Gene: GDB:THBS1; TSP1; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology; calcium binding; cell adhesion; glycoprotein; trimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248,360,708,1067/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carboxydrate (Asn) (covalent) #status absent

Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | | | | | | | |
Db 208 FQGVLQNVRFVF 219

RESULT 3
A40558
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C;Accession: A40558; A37905; B42587; S68787
R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991
A;Title: Characterization of the murine thrombospondin gene.
A;Reference number: A40558; MUID:92128941; PMID:1774063
A;Accession: A40558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1170 <LAW>
A;Cross-references: UNIPROT:P35441; UNIPARC:UPI0000028012; GB:M62449; GB:M62450; GB:M62451; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469; GB:M62470
R;Bornstein, P.; Alfifi, B.; Devarayalu, S.; Framson, P.; Li, P. J. Biol. Chem. 265, 16691-16698, 1990
A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of the 5' noncoding region
A;Reference number: A37905; MUID:90375546; PMID:2398070
A;Accession: A37905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <BOR>
A;Cross-references: UNIPARC:UPI000016D076; GB:J05605; GB:J05606; NID:G201991; PIDN:AAA404
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M. J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell cycle
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: B42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1152,'P',1154-1170 <LAH>
A;Cross-references: UNIPARC:UPI0000177A96; GB:M87276
A;Note: sequence extracted from NCBI backbone (NCBIP:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F. FEBS Lett. 387, 36-41, 1996
A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 ar
A;Reference number: S68787; MUID:96234006; PMID:8654563
A;Accession: S68787
A;Molecule type: protein
A;Residues: 19-26,'X',28-37 <CHE>
A;Cross-references: UNIPARC:UPI0000177A97
C;Complex: homotrimer, disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology; glycoprotein; homotrimer
C;Keywords: calcium binding; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | | | | | | | |
Db 208 FQGVLQNVRFVF 219

RESULT 4
TSHUP2
thrombospondin 2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: A47379; A42173
R;LaBell, T.L.; Byers, P.H. Genomics 17, 225-229, 1993
A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: poter
A;Reference number: A47379; MUID:94010892; PMID:8406456
A;Accession: A47379
A;Molecule type: mRNA
A;Residues: 1-1172 <LAB>
A;Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046680; GB:L12350; NID:G307505; PIDN:R;LaBell, T.L.; Milewicz, D.J.; Distèche, C.M.; Byers, P.H. Genomics 12, 421-429, 1992

A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression d
A:Reference number: A42173; MUID:92217961; PMID:1559694
A:Accession: A42173
A:Molecule type: mRNA
A:Residues: 560-1172 <LA2>
A:Cross-references: UNIPARC:UPI00001742C1; GB:M81339
A:Experimental source: fibroblast
A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
C:Genetics:
A:Gene: GDB:THBS2; TSP2
A:Cross-references: GDB:128789; OMIM:188061
A:Map position: 6q27-6q27
C:Complex: homotrimer, disulfide linked
C:Function:
A:Description: participates in cell migration and adhesion, and in platelet aggregation
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>
F:928-930/Region: cell attachment (R-G-D) motif
F:151,316,330,457,584,710,1069/Binding site: carboxylate (Asn) (covalent) #status predicted
F:167-226/Disulfide bonds: #status predicted
F:266,270/Disulfide bonds: interchain #status predicted
F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 68.3%; Score 41; DB 1; Length 1172;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFV 12
Db 202 FRGLQNVHLVF 213

RESULT 5
A42587
Thrombospondin 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A42587; A39851
R:Lahterty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: A42587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1172 <LAH>
A:Cross-references: UNIPROT:Q03350; UNIPARC:UPI0000029847; GB:L07803; GB:M87275; NID:g34
A:Note: sequence extracted from NCBI backbone (NCBIP:81502)
R:Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A:Reference number: A39851; MUID:91302287; PMID:1712771
A:Accession: A39851
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-873 <BOR>
A:Cross-references: UNIPARC:UPI0000160077; GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; VC
C:Keywords: calcium binding; glycoprotein
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>

Query Match 68.3%; Score 41; DB 2; Length 1172;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFV 12
Db 202 FRGLQNVHLVF 213

RESULT 6
AE2929
two component response regulator Atu3035 [imported] - Agrobacterium tumefaciens (strain C
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2929
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-747 <KUR>
A:Cross-references: UNIPROT:Q8UBI1; UNIPARC:UPI0000164787; GB:AE008689; PIDN:AAL43851.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3035
A:Map position: linear chromosome

Query Match 65.0%; Score 39; DB 2; Length 747;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGVLQNVRFV 11
Db 41 AGILSVRFV 50

RESULT 7
A98353
probable transcription regulator PA1760 [imported] - Agrobacterium tumefaciens (strain C
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A98353
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A98353
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <KUR>
A:Cross-references: UNIPROT:Q8UBI1; UNIPARC:UPI00000D2059; GB:AE007870; PIDN:AAK90347.1;
C:Genetics:
A:Gene: AGR_L_3540
A:Map position: linear chromosome

Query Match 65.0%; Score 39; DB 2; Length 783;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGVLQNVRFV 11
Db 77 AGILSVRFV 86

RESULT 8

A39804

thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A39804
R;Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <LAW>
A:Cross-references: UNIPROT:P35440; UNIPARC:UPI000013776D; GB:M60853; NID:g212763; PIDN:
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
F;325-383/Domain: von Willebrand factor type C repeat homology <WVC>
F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>

Query Match 65.0%; Score 39; DB 1; Length 1178;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FAGVLQNVRFVF 12
| | | | | : |
Db 208 FRGLLQNIHLIF 219

RESULT 9

T20985
hypothetical protein F15D4.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20985
submitted to the EMBL Data Library, September 1996

A:Reference number: Z19354

A:Accession: T20985

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-145 <WIL>

A:Cross-references: UNIPROT:Q93511; UNIPARC:UPI000007BE33; EMBL:Z80344; PIDN:CAB02486.2;

A:Experimental source: clone F15D4

C:Genetics:

A:Gene: CESP:F15D4.3

A:Map position: 2

A:Introns: 21/3; 82/1

Query Match 61.7%; Score 37; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 7.8;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FAGVLQNVRFVF 12
| | | | | : |
Db 73 FMGVAQGLRVIF 84

RESULT 10

B88349
protein F15D4.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: B88349
R;Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88349
A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-162 <STO>

A:Cross-references: UNIPROT:Q93511; UNIPARC:UPI0000179EEB; GB:chr_II; PIDN:CAB02486.1; PJ

C:Genetics:

A:Gene: F15D4.3

A:Map position: 2

Query Match 61.7%; Score 37; DB 2; Length 162;
Best Local Similarity 50.0%; Pred. No. 8.8;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FAGVLQNVRFVF 12
| | | | | : |
Db 90 FMGVAQGLRVIF 101

RESULT 11

T50315
hypothetical protein SPBC1703.02 with ARID DNA-binding domain [imported] - fission yeast
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50315
R;McDougall, R.C.; Rajandream, M.A.; Barrrell, B.G.; Cadieu, E.; Lelaure, V.; Galibert, F.

submitted to the EMBL Data Library, January 2000

A:Reference number: Z25061

A:Accession: T50315

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-780 <MCD>

A:Cross-references: UNIPROT:Q9P7W8; UNIPARC:UPI000006A13D; EMBL:AL136536; PIDN:CAB66446.1

A:Experimental source: strain 972h(-); cosmid c1703

C:Genetics:

A:Gene: SPDB:SPBC1703.02

A:Map position: 2

A:Introns: 38/2

Query Match 61.7%; Score 37; DB 2; Length 780;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGVLQNVRFVF 12
| | | | | : |
Db 93 AGILQNVYFKY 103

RESULT 12

C57253
tRNA-pseudouridine synthase (EC 5.4.99.-) - Acinetobacter calcoaceticus (fragment)
N;Alternate names: hypothetical protein lipB 5'-region
C:Species: Acinetobacter calcoaceticus
C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 05-Oct-2004
C:Accession: C57253
R;Kok, R.G.; van Thor, J.J.; Nugteren-Roodzant, I.M.; Vosman, B.; Hellingwerf, K.J.

J. Bacteriol. 177, 3295-3307, 1995

A:Title: Characterization of lipase-deficient mutants of Acinetobacter calcoaceticus BD41

A:Reference number: A57253; MUID:95286514; PMID:7768830

A:Accession: C57253

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-151 <KOK>

A:Cross-references: UNIPARC:UPI000017879A; GB:X80800

C:Superfamily: tRNA pseudouridine synthase B

C:Keywords: intramolecular transferase; isomerase; tRNA modification

Query Match 60.0%; Score 36; DB 2; Length 151;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQNVRFVF 12
| | | | | : |
Db 29 GVLQKVRWLF 38

RESULT 13

D81399
malate dehydrogenase (EC 1.1.1.37) Cj0532 [imported] - Campylobacter jejuni (strain NCTC
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81399
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-688, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A;Reference number: AB1250; MUID:20150912; PMID:10688204
A;Accession: D81399
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-300 <PAR>
A;Cross-references: UNIPROT:Q9PHY2; UNIPARC:UPI00000C216B; GB:AL111168; NID
A;Experimental source: serotype O2, strain NCTC 11168

Query Match	60.0%;	Score 36;	DB 2;	Length 440;
Best Local Similarity	63.6%;	Pred. No. 40;		
Matches	7: Conservative	2: Mismatches	2: Indels	0: Gaps

Qy 2 AGVLQNVRFVF 12
|||:|:|
pb 181 AGVLNVSAIF 191

Search completed: June 5, 2006, 22:45:04
Job time : 13.9655 secs

 α

Db 143 AGVLDNARFKY 153

|||||

RESULT 14

T50479

G protein alpha chain [imported] - Neurospora crassa

C.Species: Neurospora crassa

C.Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004

C.Accession: T50479

R.;Turner, G.E.; Borkovich, K.A.

J. Biol. Chem. 268, 14805-14811, 1993

A>Title: Identification of a G protein alpha subunit from Neurospora crassa that is a me

A:Reference number: Z25084; MUID:93315452; PMID:8325859

A:Accession: T50479

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-355 <BOR>

A:Cross-references: UNIPROT:Q05424; UNIPARC:UPI000000D1B5; EMBL:L11452; NID:g168813; PID

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; P-loop; signal transduction

Query Match 60.0%; Score 36; DB 2; Length 355;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 6: Conservative 3: Mismatches 3: Indels

Qy 1 FAGVLQNVRFVF 12
||:|:|:|
Db 75 FANILOSFRLLF 86

RESIST. 15

IAA-Ala hydrolase (IAR3) [imported] - Arabidopsis thaliana
 C Species: Arabidopsis thaliana (mouse-ear cress)
 C Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
 C Accession: F96556
 R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
 ansen, N.E.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A/A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

THIS PAGE BLANK (USPTO)


```

DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR013032; EGF_like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Cell adhesion.
FT NON_TER 1171 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 93.3%; Score 56; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred.No. 0.083; Indels 1; Gaps 0;
Matches 11; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 FAGVLQNVREVF 12
DB 207 FTGLQNVREVF 218

RESULT 2
Q28194 BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9631130; PubMed=8698834;
RA DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCP19>3.3.CO;2-0;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RA "Opposite regulation of thrombospondin-1 and corticotropin-induced
RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic
RT hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON_TER 229 229
FT NON_TER 229 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9BEC4E6B669C CRC64;

Query Match 91.7%; Score 55; DB 2; Length 229;

```

```

Best Local Similarity 91.7%; Pred.No. 0.023;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FAGVLQNVREVF 12
DB 190 FQGVQNVREVF 201

RESULT 3
Q7SY84 XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 12.
DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
TX TISSUE=Whole;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Shat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumarane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
TX TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
TX TISSUE=Whole;
RA Klein S., Strausberg R.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC054970; AAH54970.1; -; mRNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC; 1.

```

```
DR PRINTS; PR01705; TSP1RPEAT.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01208; VWFC_2; 1.
DR PROSITE; PS01184; VWFC_2; 1.
DR PROSITE; PS01184; VWFC_2; 1.
DR Cell adhesion; EGF-like domain.
KW NON TER 1
SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7BF51B CRC64;

Query Match 91.7%; Score 55; DB 2; Length 496;
Best Local Similarity 91.7%; Pred. No. 0.052;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 214 FQGVQLQNVRFVF 225

RESULT 4
Q5SPG5 BRARE
AC Q5SPG5 BRARE PRELIMINARY; PRT; 1090 AA.
ID Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.n\ (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AL928866; CA120599.1; -; Genomic DNA.
CC SMR; Q5SPG5; 751-804, 754-1089.
CC Ensembl; ENSDARG0000010785; Danio rerio.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; CoN_A like subgrp.
CC InterPro; IPR002048; EF_hand_Ca_bd.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSP_C.
CC InterPro; IPR001007; VWFC_C.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00090; TSP_1; 2.
CC Pfam; PF02412; TSP_3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC Pfam; PF00093; VWC; 1.
CC PRINTS; PR01705; TSP1RPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 2.

DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01208; VWFC_2; 1.
DR PROSITE; PS01184; VWFC_2; 1.
DR PROSITE; PS01184; VWFC_2; 1.
DR Cell adhesion; EGF-like domain.
KW NON TER 1
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1090;
Best Local Similarity 91.7%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 185 FMGVQLQNVRFVF 196

RESULT 5
TSP1 BOVIN
ID TSP1 BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
RX MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RL "cDNA cloning of bovine thrombospondin 1 and its expression in
RL odontoblasts and predentin."
RT Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta."
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3. May play a role in dentinogenesis
CC and/or maintenance of dentin and dental pulp.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- TISSUE SPECIFICITY: Odontoblasts.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB005287; BAA21115.1; -; mRNA.
CC EMBL; X87618; CAA60950.1; -; mRNA.
CC EMBL; X87619; CAA60951.1; -; mRNA.
CC PIR; S55501; S55501.
```

DR HSSP; P07996; ILSL.
 DR SMK; Q28178; 549-1169.
 DR GlycoSuiteDB; Q28178; -.
 DR InterPro; IPR013320; ConA_like_subgrp.
 DR InterPro; IPR006210; EGF.
 DR InterPro; IPR000742; EGF_3.
 DR InterPro; IPR001881; EGF_Ca_bd.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR013032; EGF_like_reg.
 DR InterPro; IPR003129; Laminin_G_TSP_N.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; TSP_3.
 DR InterPro; IPR008859; TSP_C.
 DR InterPro; IPR001007; WVF_C.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 13.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; WVC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 2.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; WVF_C; 1.
 DR PROSITE; PS01084; WVF_C; 2; 1.
 DR Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
 KW Heparin-binding; Repeat; Signal.
 FT SIGNAL 1 18 By similarity.
 FT CHAIN 19 1170 Thrombospondin-1.
 FT /FTID=PRO_0000035841.
 FT TSP N-terminal.
 FT WVF.
 FT DOMAIN 24 221 TSP type-1 1.
 FT DOMAIN 316 373 TSP type-1 2.
 FT DOMAIN 379 429 TSP type-1 3.
 FT DOMAIN 435 490 EGF-like 1.
 FT DOMAIN 549 587 EGF-like 2; calcium-binding (Potential).
 FT DOMAIN 588 645 EGF-like 3.
 FT DOMAIN 646 690 TSP type-3 1.
 FT DOMAIN 723 758 TSP type-3 2.
 FT DOMAIN 759 781 TSP type-3 3.
 FT DOMAIN 782 817 TSP type-3 4.
 FT DOMAIN 818 840 TSP type-3 5.
 FT DOMAIN 841 878 TSP type-3 6.
 FT DOMAIN 879 914 TSP type-3 7.
 FT DOMAIN 915 950 TSP C-terminal.
 FT DOMAIN 951 1170 Heparin-binding (Potential).
 FT REGION 19 232 Cell attachment site (Potential).
 FT MOTIF 926 928 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 248 248 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 360 360 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 708 708 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1085 1085 N-linked (GlcNAc...) (Potential).
 FT DISULFID 270 270 Interchain (Probable).
 FT DISULFID 274 274 Interchain (Probable).
 FT DISULFID 391 423 By similarity.
 FT DISULFID 395 428 By similarity.
 FT DISULFID 406 413 By similarity.
 FT DISULFID 447 484 By similarity.
 FT DISULFID 451 489 By similarity.
 FT DISULFID 462 474 By similarity.
 FT DISULFID 504 541 By similarity.
 FT DISULFID 508 546 By similarity.
 FT DISULFID 519 531 By similarity.
 FT DISULFID 551 562 By similarity.
 FT DISULFID 556 572 By similarity.
 FT DISULFID 575 586 By similarity.

FT DISULFID 592 608 By similarity.
 FT DISULFID 599 617 By similarity.
 FT DISULFID 620 644 By similarity.
 FT DISULFID 650 663 By similarity.
 FT DISULFID 657 676 By similarity.
 FT DISULFID 678 689 By similarity.
 FT DISULFID 705 713 By similarity.
 FT DISULFID 718 738 By similarity.
 FT DISULFID 754 774 By similarity.
 FT DISULFID 777 797 By similarity.
 FT DISULFID 813 833 By similarity.
 FT DISULFID 836 856 By similarity.
 FT DISULFID 874 894 By similarity.
 FT DISULFID 910 930 By similarity.
 FT DISULFID 946 1167 By similarity.
 FT CONFLICT 805 805 S->G (in Ref. 2).
 SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3B5FA031A CRC64;
 Query Match 91.7%; Score 55; DB 1; Length 1170;
 Best Local Similarity 91.7%; Pred. No. 0.13;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 FAGVLQNVREVF 12
 Db 208 FQGVQLQNVREVF 219
 RESULT 6
 ID TSP1_HUMAN STANDARD; PRT; 1170 AA.
 AC P07996; Q15667;
 DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
 DT 01-AUG-1988, sequence version 1.
 DT 07-MAR-2006, entry version 78.
 DE Thrombospondin-1 precursor.
 GN Name=THBS1; Synonyms=TSP, TSP1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Endothelial cell;
 MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
 RX Lawler J., Hynes R.O.;
 RA "The structure of human thrombospondin, an adhesive glycoprotein with
 RT multiple calcium-binding sites and homologies with several different
 RT proteins.";
 RL J. Cell Biol. 103:1635-1648(1986).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
 RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
 RA Baumgartel D.M., Rotwein P., Frazier W.A.;
 RT "Complete thrombospondin mRNA sequence includes potential regulatory
 RT sites in the 3' untranslated region.";
 RL J. Cell Biol. 108:729-736(1989).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 1-397.
 RX MEDLINE=87157592; PubMed=3030396;
 RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
 RA "Partial amino acid sequence of human thrombospondin as determined by
 RT analysis of cDNA clones: homology to malarial circumsporozoite
 RT proteins.";
 RL Biochemistry 25:8418-8425(1986).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 1-374.
 RX MEDLINE=86287276; PubMed=3461443;
 RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
 RA "Characterization of a cDNA encoding the heparin and collagen binding
 RT domains of human thrombospondin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).

[5]
RN NP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RL gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
[6]
RN NP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA La Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RL from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
[7]
RN NP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;
RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RL module.";
RL J. Biol. Chem. 276:6485-6498(2001).
[8]
RN NP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annie D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RL of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
[9]
RN NP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RL hydrazide chemistry, and mass spectrometry.";
RL J. Proteome Res 4:2070-2080(2005).
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC beta-3 and alpha-IIB/beta-3.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 WFCC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04665; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAA61237.1; ALT SEQ; mRNA.
DR EMBL; J04835; AAA61178.1; -; Genomic DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUPL.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1Z44; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; gene.
DR Reactome; P07996; -.

DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF_3.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
DR 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Calcium; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTid=PRO_0000035842.
FT TSP N-terminal.
FT VWF_1.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT C-linked (Man).
FT /FTid=CAR_000205.
FT O-linked (Fuc...).
FT /FTid=CAR_000206.
FT C-linked (Man).
FT /FTid=CAR_000207.
FT C-linked (Man).
FT /FTid=CAR_000208.
FT O-linked (Fuc...).
FT /FTid=CAR_000209.
FT C-linked (Man).
FT /FTid=CAR_000210.
FT O-linked (Fuc...).
FT /FTid=CAR_000211.
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...).
FT CARBOHYD 394 394
FT CARBOHYD 438 438
FT CARBOHYD 441 441
FT CARBOHYD 450 450
FT CARBOHYD 498 498
FT CARBOHYD 507 507
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067

FT DISULFID 270 270 Interchain (Probable).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546

Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FAGVLQNVRFVF 12

Db 208 FQGVQNVRFVF 219

RESULT 7

TSPI_MOUSE
ID TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 57.
DE Thrombospondin-1 precursor.
GN Name=Thbel; Synonyms=Tspl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
[2]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
[3]
RN NUCLEOTIDE SEQUENCE OF 1-490.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Ali D., Devarayalu S., Framson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
[4]
RN PROTEIN SEQUENCE OF 19-37.
RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
RT "Expression and initial characterization of recombinant mouse thrombospondin 1 and initial characterization of recombinant mouse thrombospondin 3.";
RL FEBS Lett. 387:36-41(1996).
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibrinogen, fibronectin, laminin, type V collagen and integrins alpha-V/beta-1, alpha-3/beta-3 and alpha-1Ib/beta-3.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.

CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWF domain.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; M62470; AAA50611.1; -; Genomic DNA.
DR EMBL; M62450; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62451; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62452; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62453; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62454; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62455; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62456; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62457; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62458; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62459; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62460; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62461; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62462; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62463; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62464; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62465; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62466; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62467; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62468; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62469; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M87276; AAA53063.1; -; mRNA.
DR EMBL; J05606; AAA40431.1; -; Genomic DNA.
DR EMBL; J05605; AAA40431.1; -; Genomic DNA.
DR PIR; A40558; A40558.
DR HSR; P07996; 1L5L.
DR SWR; P35441; 549-1169.
DR Ensembl; ENSMUSG0000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR003367; TSP3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR01007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP1; 3.
DR Pfam; PF02412; TSP3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS01208; VWF_C; 1.
KW Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
KW Glycoprotein; Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18 Thrombospondin-1.
FT CHAIN 19 1170 /FTID=PRO_0000035843.
FT DOMAIN 24 221 TSP N-terminal.
FT DOMAIN 316 373 VWF.
FT DOMAIN 379 429 TSP type-1 1.

FT	DOMAIN	435	490	TSP type-1 2.
FT	DOMAIN	492	547	TSP type-1 3.
FT	DOMAIN	549	587	EGF-like 1.
FT	DOMAIN	588	645	EGF-like 2; calcium-binding (Potential).
FT	DOMAIN	646	690	EGF-like 3.
FT	DOMAIN	723	758	TSP type-3 1.
FT	DOMAIN	759	781	TSP type-3 2.
FT	DOMAIN	782	817	TSP type-3 3.
FT	DOMAIN	818	840	TSP type-3 4.
FT	DOMAIN	841	878	TSP type-3 5.
FT	DOMAIN	879	914	TSP type-3 6.
FT	DOMAIN	915	950	TSP type-3 7.
FT	DOMAIN	951	1170	TSP C-terminal.
FT	REGION	19	232	Heparin-binding (Potential).
FT	MOTIF	926	928	Cell attachment site (Potential).
FT	CARBOHYD	248	248	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	360	360	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	708	708	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1067	1067	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	270	270	Interchain (Probable).
FT	DISULFID	274	274	Interchain (Probable).
FT	DISULFID	391	423	By similarity.
FT	DISULFID	395	428	By similarity.
FT	DISULFID	406	413	By similarity.
FT	DISULFID	447	484	By similarity.
FT	DISULFID	451	489	By similarity.
FT	DISULFID	462	474	By similarity.
FT	DISULFID	504	541	By similarity.
FT	DISULFID	508	546	By similarity.
FT	DISULFID	519	531	By similarity.
FT	DISULFID	551	562	By similarity.
FT	DISULFID	556	572	By similarity.
FT	DISULFID	575	586	By similarity.
FT	DISULFID	592	608	By similarity.
FT	DISULFID	599	617	By similarity.
FT	DISULFID	620	644	By similarity.
FT	DISULFID	650	663	By similarity.
FT	DISULFID	657	676	By similarity.
FT	DISULFID	678	689	By similarity.
FT	DISULFID	705	713	By similarity.
FT	DISULFID	718	738	By similarity.
FT	DISULFID	754	774	By similarity.
FT	DISULFID	777	797	By similarity.
FT	DISULFID	833	833	By similarity.
FT	DISULFID	836	856	By similarity.
FT	DISULFID	874	894	By similarity.
FT	DISULFID	910	930	By similarity.
FT	DISULFID	946	1167	By similarity.
FT	CONFLICT	1025	1025	F -> L (in Ref. 2).
SQ	SEQUENCE	1170 AA; 129647 MW; 0443E493615E7F06 CRC64;		
Query Match				
Best Local Similarity 91.7%; Score 55; DB 1; Length 1170;				
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	FAGYLQNVRFVF 12		
Db	208	FOGVLQNVRFVF 219		
RESULT 8				
Q3TR40	MOUSE	PRELIMINARY;	PRT; 1170 AA.	
ID	Q3TR40			
AC	Q3TR40			
DT	11-OCT-2005,	integrated into UniProtKB/TrEMBL.		
DT	11-OCT-2005,	sequence version 1.		
DE	07-FEB-2006,	entry version 5.		
DE	Adult male aorta and vein cDNA, RIKEN full-length enriched library,			
DE	clone:A530055N06 product:thrombospondin 1, full insert sequence.			
GN	Name=thbs1;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			

OC	Muroidea; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA	Carninci P., Hayashizaki Y.;
RA	"High-efficiency full-length cDNA cloning.";
RL	Methods Enzymol. 303:19-44(1999).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX	PubMed=16141072; DOI=10.1126/science.1112014;
RA	Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA	Oyama K., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA	Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA	Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA	Ambsel-Imbimbato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA	Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA	Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA	Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA	di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA	Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA	Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA	Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA	Hill D., Humnicki L., Iacono M., Ikeo K., Iwano A., Ishikawa T.,
RA	Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA	Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA	Kurochkin I.V., Laureau L.F., Lazarevic D., Lipovich L., Liu J.,
RA	Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA	Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA	Mottagui-Fabriz S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA	Nillesen R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA	Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA	Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA	Rost B., Ruan Y., Salzberg S.L., Sadelin A., Schneider C.,
RA	Schombach C., Sekiguchi K., Semple C.A., Seno S., Sesca L., Sheng Y.,
RA	Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA	Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
RA	Tamofa K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA	Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA	Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA	Grimmond S.M., Teasdale R.D., Liu E.T., Bruscia V., Quackenbush J.,
RA	Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA	Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA	Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA	Kawashima T., Kojima M., Kondo S., Konno K., Nakano K., Ninomiya N.,
RA	Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA	Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA	Hayashizaki Y.;
RA	"The transcriptional landscape of the mammalian genome.";
RL	Science 309:1559-1563(2005).
RN	[3]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX	PubMed=16141073; DOI=10.1126/science.1112009;
RG	RIKEN Genome Exploration Research Group, and Genome Science Group
RG	(Genome Network Core Team) and the FANTOM Consortium;
RT	"Antisense Transcription in the Mammalian Transcriptome.";
RL	Science 309:1564-1566(2005).
RN	[4]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA	Blake J.A., Bradt D., Bruscia V., Chothia K., Corbani L.E., Cousins S.,
RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer J.K.S.,
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,


```

CC EMBL; AF309630; AAQ14549.1; -; mRNA.
DR SML; Q7ISA3; 834-1169.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tgp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE; PS0184; VWFC_2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFV 12
Db 208 FQGVQLQNVRFV 219

RESULT 10
Q80YQ1_MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thb1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX Director MGC Project;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC050917; AAH50917.1; -; mRNA.
DR HSSP; P07996; ILSL.
DR SMR; Q80YQ1; 835-1170.
DR Ensembl; ENSMUSG00000040152; Mus musculus.
DR MGI; MGI:98737; Thb1.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tgp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE; PS0184; VWFC_2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFV 12
Db 208 FQGVQLQNVRFV 219

RESULT 11
Q8CGB2_MOUSE PRELIMINARY; PRT; 1171 AA.
ID Q8CGB2_MOUSE
AC Q8CGB2;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-

```

length enriched library, clone:G930018021 product:thrombospondin 1,
full insert sequence).
GN Name=Thb1;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=C2ECH II; TISSUB=Mammary tumor metastatized to lung. Tumor
RC arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Colling P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.M., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kryzhanzki M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=C2ECH II; TISSUB=Mammary tumor metastatized to lung. Tumor
RC arose spontaneously;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning.";
RT Methods Enzymol. 303:19-44 (1999).
[4]
RP NUCLEOTIDE SEQUENCE
RC TISSUB=Mammary gland;
RX PubMed=16141072; DOI=10.1126/science.11112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bono H., Chalk A.M.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottaqui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,

Schombach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiyama K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K., Bult C.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563 (2005).
[5]
RP NUCLEOTIDE SEQUENCE
RC TISSUB=Mammary gland;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RA RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566 (2005).
[6]
RP NUCLEOTIDE SEQUENCE
RC TISSUB=Mammary gland;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Negashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
[7]
RP NUCLEOTIDE SEQUENCE
RC TISSUB=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaehizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RN (8)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayaehizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN (9)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayaehizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN (10)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayaehida K.,
RA Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayaehizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: BC042422; AAH42422.1; -: mRNA.
CC EMBL: AK145202; BAE26293.1; -: mRNA.
CC HSSP: P07996; 1LSL.

Query Match 91.7%; Score 55; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. NO. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PAGVLQNRVFFV 12
Db 208 FQGLVQNRVFFV 219

RESULT 12
TSPI_XENLA STANDARD; PRT: 1173 AA.
AC P35448;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 54.
DE Thrombospondin-1 precursor.
GN Name=thbs1; Synonyms=tspl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN (1)

RP NUCLEOTIDE SEQUENCE [MRNA].
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RT "Cloning, characterization and expression of thrombospondin-1 in
RT Xenopus laevis embryos.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: L04278; -: NOT_ANNOTATED_CDS; mRNA.
CC HSSP: P07996; 1LSL.
CC SMR: P35448; 552-1172.
DR InterPro: IPR013320; ConA_like_subgrp.
DR InterPro: IPR006210; EGF.
DR InterPro: IPR000742; EGF_3.
DR InterPro: IPR001881; EGF_Ca_bd.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR013032; EGF_like_reg.
DR InterPro: IPR003129; Laminin_G_TSP_N.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR008085; TSP_1.
DR InterPro: IPR003367; tsp_3.
DR InterPro: IPR008859; TSP_C.
DR InterPro: IPR001007; VWFC_C.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00181; EGF_2.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS50026; EGF_3; 2.
DR PROSITE: PS50092; TSP1; 3.
DR PROSITE: PS01208; VWFC_1; 1.
DR PROSITE: PS0184; VWFC_2; 1.
DR Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1173 Thrombospondin-1.
FT TSP N-terminal.
FT DOMAIN 23 224 VWFC.
FT DOMAIN 319 376 TSP type-1 1.
FT DOMAIN 382 432 TSP type-1 2.
FT DOMAIN 438 493 TSP type-1 3.
FT DOMAIN 495 550 EGF-like 1.
FT DOMAIN 550 590 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 591 648 EGF-like 3.
FT DOMAIN 649 693 TSP type-3 1.
FT DOMAIN 726 761 TSP type-3 2.
FT DOMAIN 762 784 TSP type-3 3.
FT DOMAIN 785 820 TSP type-3 4.
FT DOMAIN 821 843 TSP type-3 5.
FT DOMAIN 844 881 TSP type-3 6.
FT DOMAIN 882 917 TSP type-3 7.
FT DOMAIN 918 953 TSP C-terminal.
FT DOMAIN 954 1173

FT REGION 23 235 Heparin-binding (Potential).
 FT MOTIF 929 931 Cell attachment site (Potential).
 FT CARBOHYD 155 155 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 158 158 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 250 250 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 363 363 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 705 705 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 711 711 N-linked (GlcNAc . . .) (Potential).
 FT CARBOHYD 1070 1070 N-linked (GlcNAc . . .) (Potential).
 FT DISULFID 394 426 By similarity.
 FT DISULFID 398 431 By similarity.
 FT DISULFID 409 416 By similarity.
 FT DISULFID 450 487 By similarity.
 FT DISULFID 454 492 By similarity.
 FT DISULFID 465 477 By similarity.
 FT DISULFID 507 544 By similarity.
 FT DISULFID 511 549 By similarity.
 FT DISULFID 522 534 By similarity.
 FT DISULFID 554 565 By similarity.
 FT DISULFID 559 575 By similarity.
 FT DISULFID 578 589 By similarity.
 FT DISULFID 595 611 By similarity.
 FT DISULFID 602 620 By similarity.
 FT DISULFID 623 647 By similarity.
 FT DISULFID 653 666 By similarity.
 FT DISULFID 660 679 By similarity.
 FT DISULFID 681 692 By similarity.
 FT DISULFID 708 716 By similarity.
 FT DISULFID 721 741 By similarity.
 FT DISULFID 757 777 By similarity.
 FT DISULFID 780 800 By similarity.
 FT DISULFID 816 836 By similarity.
 FT DISULFID 839 859 By similarity.
 FT DISULFID 877 897 By similarity.
 FT DISULFID 913 933 By similarity.
 FT DISULFID 949 1170 By similarity.
 SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 91.7%; Score 55; DB 1; Length 1173;
 Best Local Similarity 91.7%; Pred. No. 0.13;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVREVF 12
 | | | | | | | | | | | | | | | |
 Db 211 FQGVQLQNVREVF 222

RESULT 13

Q4S758 TETNG
 ID Q4S758 TETNG PRELIMINARY; PRT; 1193 AA.
 AC Q4S758;
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DT 21-FEB-2006, entry version 8.
 DE Chromosome 14 SCAP14723, whole genome shotgun sequence. (Fragment).
 GN ORFNames=GSTENG0022976001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OC NCBI_TaxID=99883;
 RN [1]

NUCLEOTIDE SEQUENCE

RP PubMed=15496914; DOI=10.1038/nature03025;
 RX Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RL the early vertebrate proto-karyotype.";
 RN Nature 431:946-957(2004).
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; CAAE01014723; CAG03524.1; -; Genomic_DNA.
 DR SMR; Q4S758; 811-1148.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0008201; F:heparin binding; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR006210; EGF.
 DR InterPro; IPR000742; EGF_3.
 DR InterPro; IPR001881; EGF_Ca bd.
 DR InterPro; IPR013032; EGF_like_reg.
 DR InterPro; IPR003129; Laminin_G_TSP_N.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP 1.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 12.
 DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; WVC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS01208; VWF_C_1; 1.
 DR PROSITE; PS50184; VWF_C_2; 1.
 KW Cell adhesion.
 FT NON TER 1 1
 SQ SEQUENCE 1193 AA; 133256 MW; 6B8781648FCEC7F2 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1193;
 Best Local Similarity 91.7%; Pred. No. 0.13;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVREVF 12
 | | | | | | | | | | | | | | | |
 Db 187 FMGVQLQNVREVF 198

RESULT 14

Q59E99 HUMAN
 ID Q59E99 HUMAN PRELIMINARY; PRT; 1225 AA.
 AC Q59E99;
 DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
 DT 26-APR-2005, sequence version 1.
 DT 21-FEB-2006, entry version 10.
 DE Thrombospondin 1 variant (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Rutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

```
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Toki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
DR EMBL; AB209912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939; 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_2; 1.
DR PROSITE; PS0184; VWF_2; 1.
DR Cell adhesion; EGF-like domain.
KW NON_TER
FT NON_TER
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16E57157B12 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1225;
Best Local Similarity 91.7%; Pred. NO. 0.14;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 263 FQGLQNVRFVF 274

RESULT 15
Q5U903_PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1; (Fig).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OX Sub.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang K., Maucó G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
DR EMBL; AY773342; AAV38110.1; -; mRNA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_2; 1.
DR NON_TER
FT NON_TER
SQ SEQUENCE 249 AA; 27560 MW; 465D664BB0329C32 CRC64;

Query Match 83.3%; Score 50; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. NO. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GVLQNVRFVF 12
Db 2 GVLQNVRFVF 11

Search completed: June 5, 2006, 22:42:46
Job time : 109.931 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-26

Perfect score: 60

Sequence: 1 FAGVLQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5 COMB pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6 COMB pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7 COMB pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H COMB pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS COMB pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE COMB pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	91.7	825	2	US-09-949-002-482
2	55	91.7	831	2	US-09-939-853A-97
3	55	91.7	831	2	US-09-939-853A-98
4	55	91.7	1170	1	US-08-313-288B-20
5	55	91.7	1170	2	US-09-657-472-2
6	55	91.7	1170	2	US-09-949-002-350
7	41	68.3	1045	2	US-09-949-016-11112
8	41	68.3	1172	1	US-08-313-288B-19
9	41	68.3	1172	2	US-09-949-016-6333
10	38	63.3	740	2	US-09-134-000C-6441
11	37	61.7	2475	2	US-09-413-814-48
12	36	60.0	61	2	US-09-270-767-61396
13	36	60.0	304	2	US-09-270-767-45864
14	36	60.0	335	2	US-09-270-767-4451
15	36	60.0	476	2	US-09-489-039A-8120
16	35	58.3	345	2	US-09-252-991A-28350
17	34	56.7	89	2	US-09-270-767-46041
18	34	56.7	204	2	US-09-902-540-16431
19	34	56.7	436	2	US-09-252-991A-28888
20	34	56.7	439	1	US-08-507-431-4
21	34	56.7	439	1	US-08-700-546-2
22	34	56.7	439	1	US-08-902-655A-4
23	34	56.7	439	2	US-09-116-622-4
24	34	56.7	439	2	US-09-219-277-4
25	34	56.7	439	2	US-09-599-661-4
26	34	56.7	439	2	US-09-543-681A-8164

```

27 34 56.7 479 2 US-09-328-352-4361 Sequence 4361, Ap
28 34 56.7 541 2 US-08-687-590-28 Sequence 28, Appl
29 34 56.7 541 2 US-09-311-311C-25 Sequence 25, Appl
30 34 56.7 715 1 US-08-484-993B-10 Sequence 10, Appl
31 34 56.7 715 1 US-08-484-158B-10 Sequence 10, Appl
32 34 56.7 715 1 US-08-484-596A-10 Sequence 10, Appl
33 34 56.7 715 1 US-08-480-150A-10 Sequence 10, Appl
34 34 56.7 715 2 US-08-458-731-10 Sequence 10, Appl
35 34 56.7 715 2 US-08-149-223A-10 Sequence 10, Appl
36 34 56.7 731 2 US-09-270-767-42057 Sequence 10608, A
37 34 56.7 749 2 US-09-902-540-10608 Sequence 3, Appli
38 33 55.0 40 1 US-07-868-353A-3 Sequence 3, Appli
39 33 55.0 40 1 US-08-407-804-3 Sequence 3, Appli
40 33 55.0 40 2 US-09-124-807-3 Sequence 3, Appli
41 33 55.0 135 2 US-09-270-767-33170 Sequence 33170, A
42 33 55.0 135 2 US-09-270-767-48387 Sequence 48387, A
43 33 55.0 149 2 US-09-270-767-42044 Sequence 42044, A
44 33 55.0 273 2 US-09-710-279-84 Sequence 84, Appl
45 33 55.0 275 2 US-09-134-001C-3732 Sequence 3732, Ap

```

ALIGNMENTS

```

RESULT 1
US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-482

```

```

Query Match          91.7%; Score 55; DB 2; Length 825;
Best Local Similarity 91.7%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 FAGVLQNVRFVF 12
      |||||
Db      308 FQGVQLQNVRFVF 319

```

```

RESULT 2
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20

```

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.7%; Score 55; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          91.7%; Score 55; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.034;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.034;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```



```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-350

Query Match          91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.034;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
      | | | | | | | | | |
Db      208 FQGVQLNVRFVF 219

RESULT 7
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11112

Query Match          68.3%; Score 41; DB 2; Length 1045;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
      | | | | | | | |
Db      281 FRGLLQNVHLVF 292

RESULT 8
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
```

```
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match          68.3%; Score 41; DB 1; Length 1172;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
      | | | | | | | |
Db      202 FRGLLQNVHLVF 213

RESULT 9
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6333

Query Match          68.3%; Score 41; DB 2; Length 1172;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
      | | | | | | | |
Db      202 FRGLLQNVHLVF 213

RESULT 10
```

```

US-09-134-000C-6441
; Sequence 6441, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6441
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (11)..(11)
; OTHER INFORMATION: Amino acid 11 is Xaa wherein Xaa = any amino acid.
US-09-134-000C-6441

Query Match          63.3%; Score 38; DB 2; Length 740;
Best Local Similarity 41.7%; Pred. No. 47;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0

```

Qy	1 FAGVLQNVRFFV 12
	: :
Db	728 FCGILGNINFIY 739

```

RESULT 11
US-09-413-814-48
; Sequence 48, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 2475
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-48

```

Query Match	61.7%	Score 37;	DB 2;	Length 2475;
Best Local Similarity	58.3%;	pred. NO. 2.8e+02;		
Matches	7;	Conservative	2;	Mismatches 3;
				Indels 0;
				Gaps 0;

Qy	1 FAGVLQNVRVF 12
	:
Dδ	123 FAGVSSNLSFLF 134

```

RESULT 12
US-09-270-767-61396
; Sequence 61396, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61396
; LENGTH: 61
; TYPE: PR1
; ORGANISM: Drosophila melanogaster
US-09-270-767-61396

Query Match 60.0%; Score 36; DB 2; Length 61;
Best Local Similarity 58.3%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps

```

Qy	1	FAGVLQNRFEV	12
Db	44	FGGVTNVRLLF	55

RESULT 13
 US-09-270-767-45864
 ; Sequence 45864, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 45864
 ; LENGTH: 304
 ; TYPE: PRT
 ; ORGANISM: *Drosophila melanogaster*
 ; FEATURE:
 ; OTHER INFORMATION: Xaa means any amino acid
 US-09-270-767-45864

Query Match	60.0%	Score 36;	DB 2;	Length 304;
Best Local Similarity	58.3%	Pred. No. 44;		
Matches	7;	Conservative	2;	Mismatches
				3; Indels 0; Gaps 0;

Qy 1 FAGVLQNRVVF 12
| | | | | | | |
| | | | | | | |
Dp 287 FFGVTVNRLLF 298

```

RESULT 14
US-09-270-767-46451
; Sequence 46451, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.;
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46451
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:

```

; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-46451

Query Match 60.0%; Score 36; DB 2; Length 335;
Best Local Similarity 58.3%; Pred. NO. 48;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PAGVLQNVRFVF 12
| | | | | | | |
Db 159 PFGXLQNRRLIF 170

RESULT 15

US-09-489-039A-8120
; Sequence 8120, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8120
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8120

Query Match 60.0%; Score 36; DB 2; Length 476;
Best Local Similarity 54.5%; Pred. NO. 71;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AGVLQNVRFVF 12
| | | | | | | |
Db 33 AGTLKNMRYAF 43

Search completed: June 5, 2006, 22:48:57
Job time : 24.8966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-26

Perfect score: 60

Sequence: 1 FAGVLQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	91.7	12	4	US-10-474-213-28
2	55	91.7	240	4	Sequence 28, Appl
3	55	91.7	240	5	Sequence 40, Appl
4	55	91.7	432	5	Sequence 40, Appl
5	55	91.7	432	5	Sequence 1020, Ap
6	55	91.7	459	6	Sequence 1022, Ap
7	55	91.7	466	3	Sequence 462, App
8	55	91.7	555	6	Sequence 1047, Ap
9	55	91.7	578	6	Sequence 454, App
10	55	91.7	685	6	Sequence 456, App
11	55	91.7	804	6	Sequence 452, App
12	55	91.7	828	6	Sequence 453, App
13	55	91.7	831	3	Sequence 455, App
14	55	91.7	831	3	Sequence 97, Appl
15	55	91.7	855	6	Sequence 98, Appl
16	55	91.7	1000	6	Sequence 461, App
17	55	91.7	1105	6	Sequence 457, App
18	55	91.7	1150	6	Sequence 458, App
19	55	91.7	1152	3	Sequence 1, Appli
20	55	91.7	1169	5	Sequence 1, Appli
21	55	91.7	1170	4	Sequence 7, Appli
22	55	91.7	1170	4	Sequence 12, Appl
23	55	91.7	1170	4	Sequence 2, Appli
24	55	91.7	1170	4	Sequence 114, App
25	55	91.7	1170	4	Sequence 2, Appli
26	55	91.7	1170	4	Sequence 1170, Ap
27	55	91.7	1170	4	Sequence 38, Appl
					Sequence 482, App

28 55 91.7 1170 4 US-10-419-462-38 Sequence 38, Appl
29 55 91.7 1170 5 US-10-741-600-1018 Sequence 1018, Ap
30 55 91.7 1170 5 US-10-741-600-1019 Sequence 1019, Ap
31 55 91.7 1170 5 US-10-741-600-1021 Sequence 1021, Ap
32 55 91.7 1170 5 US-10-782-968-38 Sequence 38, Appl
33 55 91.7 1170 5 US-10-849-989-44 Sequence 44, Appl
34 55 91.7 1170 5 US-10-631-467-548 Sequence 548, App
35 55 91.7 1170 5 US-10-631-467-1376 Sequence 1376, Ap
36 55 91.7 1170 5 US-10-831-997-2 Sequence 2, Appli
37 55 91.7 1170 5 US-10-995-561-594 Sequence 594, App
38 55 91.7 1170 5 US-10-995-561-595 Sequence 595, App
39 55 91.7 1170 5 US-10-995-561-596 Sequence 596, App
40 55 91.7 1170 6 US-11-037-713-51 Sequence 51, Appl
41 55 91.7 1170 6 US-11-046-644-28 Sequence 28, Appl
42 55 91.7 1170 6 US-11-046-456-28 Sequence 28, Appl
43 41 68.3 16 3 US-09-822-682-6 Sequence 6, Appli
44 41 68.3 16 5 US-10-997-699-6 Sequence 6, Appli
45 41 68.3 151 3 US-09-925-299-1251 Sequence 1251, Ap

ALIGNMENTS

RESULT 1

US-10-474-213-28

; Sequence 28, Application US/10474213

; Publication No. US20040214248A1

; GENERAL INFORMATION:

; APPLICANT: Roberts, David D

; APPLICANT: Krutzsch, Henry C

; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF

; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 224329

; CURRENT APPLICATION NUMBER: US/10/474,213

; CURRENT FILING DATE: 2003-10-06

; PRIOR APPLICATION NUMBER: PCT/US02/10535

; PRIOR FILING DATE: 2002-04-03

; PRIOR APPLICATION NUMBER: 60/281,994

; PRIOR FILING DATE: 2001-04-06

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 28

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic peptide

US-10-474-213-28

Query Match 91.7%; Score 55; DB 4; Length 12;

Best Local Similarity 91.7%; Pred. No. 0.00077;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 FAGVLQNVRFVF 12

1 FQGLQNVRFVF 12

Db

RESULT 2

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Williams, Kevin J.

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 40

```
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match          91.7%; Score 55; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 190 FQGVQLQNVRFVF 201

RESULT 3
US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match          91.7%; Score 55; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 190 FQGVQLQNVRFVF 201

RESULT 4
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match          91.7%; Score 55; DB 5; Length 432;
```

```
Best Local Similarity 91.7%; Pred. No. 0.042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 5
US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match          91.7%; Score 55; DB 5; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 6
US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match          91.7%; Score 55; DB 6; Length 459;
Best Local Similarity 91.7%; Pred. No. 0.045;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
Db 208 FQGVQLQNVRFVF 219

RESULT 7
US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 91.7%; Score 55; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.046;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | |
Db 261 FQGVQLQNVRFVF 272

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 91.7%; Score 55; DB 6; Length 555;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQLQNVRFVF 219

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 91.7%; Score 55; DB 6; Length 578;
Best Local Similarity 91.7%; Pred. No. 0.058;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQLQNVRFVF 219

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 91.7%; Score 55; DB 6; Length 685;
Best Local Similarity 91.7%; Pred. No. 0.071;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQLQNVRFVF 219

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 91.7%; Score 55; DB 6; Length 804;
Best Local Similarity 91.7%; Pred. No. 0.084;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FAGVLQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQLQNVRFVF 219

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match          91.7%; Score 55; DB 6; Length 828;
Best Local Similarity 91.7%; Pred. No. 0.087;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
Db      208 FQGVQLQNVRFVF 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.7%; Score 55; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.088;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
Db      208 FQGVQLQNVRFVF 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-09-939-853A-98

Query Match          91.7%; Score 55; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.088;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
Db      208 FQGVQLQNVRFVF 219

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match          91.7%; Score 55; DB 6; Length 855;
Best Local Similarity 91.7%; Pred. No. 0.09;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FAGVLQNVRFVF 12
Db      208 FQGVQLQNVRFVF 219

Search completed: June 6, 2006, 00:00:11
Job time : 78.6207 secs
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-26

Perfect score: 60

Sequence: 1 FAGVLQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US09_NEW_PUB.pap:*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US06_NEW_PUB.pap:*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US07_NEW_PUB.pap:*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US08_NEW_PUB.pap:*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/PCT_NEW_PUB.pap:*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US10_NEW_PUB.pap:*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US11_NEW_PUB.pap:*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	60.0	440	US-10-953-349-1348	Sequence 1348, Ap
2	34	56.7	182	US-10-953-349-15318	Sequence 15318, A
3	33	55.0	290	US-10-953-349-5462	Sequence 5462, Ap
4	33	55.0	329	US-10-953-349-5461	Sequence 5461, Ap
5	33	55.0	608	US-11-203-828-5	Sequence 5, Appli
6	32	53.3	297	US-11-293-697-4132	Sequence 4132, Ap
7	32	53.3	570	US-10-953-349-32501	Sequence 32501, A
8	32	53.3	577	US-10-953-349-32500	Sequence 32500, A
9	32	53.3	581	US-10-953-349-32499	Sequence 32499, A
10	31	51.7	298	US-10-953-349-9347	Sequence 9347, Ap
11	31	51.7	1234	US-10-953-349-3997	Sequence 3997, Ap
12	31	51.7	1237	US-10-953-349-3996	Sequence 3996, Ap
13	31	51.7	1337	US-10-953-349-3995	Sequence 3995, Ap
14	30	50.0	74	US-10-953-349-15283	Sequence 15283, A
15	30	50.0	77	US-10-953-349-15282	Sequence 15282, A
16	30	50.0	227	US-10-953-349-39138	Sequence 39138, A
17	30	50.0	265	US-10-953-349-39137	Sequence 39137, A
18	30	50.0	276	US-10-953-349-39136	Sequence 39136, A
19	30	50.0	386	US-10-953-349-23720	Sequence 23720, A
20	30	50.0	459	US-10-953-349-7070	Sequence 7070, Ap
21	30	50.0	468	US-10-953-349-2644	Sequence 2644, Ap
22	30	50.0	470	US-10-953-349-7069	Sequence 7069, Ap
23	30	50.0	479	US-10-953-349-2643	Sequence 2643, Ap
24	30	50.0	480	US-10-953-349-23719	Sequence 23719, A
25	30	50.0	489	US-10-953-349-2642	Sequence 2642, Ap

26	30	50.0	521	6	US-10-953-349-12464	Sequence 12464, A
27	30	50.0	528	6	US-10-953-349-23718	Sequence 23718, A
28	30	50.0	534	6	US-10-953-349-12463	Sequence 12463, A
29	30	50.0	557	6	US-10-953-349-12462	Sequence 12462, A
30	30	50.0	750	6	US-10-511-937-2413	Sequence 2413, Ap
31	30	50.0	843	6	US-10-953-349-16527	Sequence 16527, A
32	30	50.0	855	7	US-11-247-437-2	Sequence 2, Appli
33	30	50.0	962	6	US-10-953-349-16526	Sequence 16526, A
34	30	50.0	1014	6	US-10-953-349-16525	Sequence 16525, A
35	30	50.0	1043	6	US-10-511-937-2452	Sequence 2452, Ap
36	30	50.0	1504	6	US-10-505-928-662	Sequence 662, App
37	29	48.3	139	7	US-11-293-697-3323	Sequence 3323, Ap
38	29	48.3	164	6	US-10-953-349-1159	Sequence 1159, Ap
39	29	48.3	186	6	US-10-953-349-27518	Sequence 27518, A
40	29	48.3	188	6	US-10-953-349-27517	Sequence 27517, A
41	29	48.3	206	6	US-10-953-349-12003	Sequence 12003, A
42	29	48.3	206	6	US-10-953-349-20955	Sequence 20955, A
43	29	48.3	213	6	US-10-953-349-27516	Sequence 27516, A
44	29	48.3	249	6	US-10-953-349-1158	Sequence 1158, Ap
45	29	48.3	249	6	US-10-953-349-11050	Sequence 11050, A

ALIGNMENTS

RESULT 1
US-10-953-349-1348
; Sequence 1348, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1348
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1348

Query Match 60.0%; Score 36; DB 6; Length 440;
Best Local Similarity 63.6%; Pred. No. 9.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AGVLQNVRFVF 12
|||:|:|:|
Db 181 AGVLENVSAIF 191

RESULT 2
US-10-953-349-15318
; Sequence 15318, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15318
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15318

Query Match 56.7%; Score 34; DB 6; Length 182;


```
Qy      2 AGVLQVRVFV 11
Db      414 AGVLNVQVYM 423

RESULT 8
US-10-953-349-32500
; Sequence 32500, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32500
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-32500

Query Match      53.3%; Score 32; DB 6; Length 577;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AGVLQVRVFV 11
Db      421 AGVLNVQVYM 430

RESULT 9
US-10-953-349-32499
; Sequence 32499, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32499
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-32499

Query Match      53.3%; Score 32; DB 6; Length 581;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AGVLQVRVFV 11
Db      425 AGVLNVQVYM 434

RESULT 10
US-10-953-349-9347
; Sequence 9347, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
```

```
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9347

Query Match      51.7%; Score 31; DB 6; Length 298;
Best Local Similarity 44.4%; Pred. No. 56;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      3 GVLQNVRFV 11
Db      88 GILQGVRFI 96

RESULT 11
US-10-953-349-3997
; Sequence 3997, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3997
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3997

Query Match      51.7%; Score 31; DB 6; Length 1234;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VLQNVRFV 12
Db      1122 ILENVRLVY 1130

RESULT 12
US-10-953-349-3996
; Sequence 3996, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3996
; LENGTH: 1237
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3996

Query Match      51.7%; Score 31; DB 6; Length 1237;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VLQNVRFV 12
Db      1125 ILENVRLVY 1133
```

```
RESULT 13
US-10-953-349-3995
; Sequence 3995, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3995
; LENGTH: 1337
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3995

Query Match          51.7%; Score 31; DB 6; Length 1337;
Best Local Similarity 55.6%; Pred. No. 2.9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VLQNVREVF 12
Db      1225 ILENRLVY 1233

RESULT 14
US-10-953-349-15283
; Sequence 15283, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15283
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15283

Query Match          50.0%; Score 30; DB 6; Length 74;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      4 VLQNVREVF 12
Db      21 IVQGVRFAP 29

RESULT 15
US-10-953-349-15282
; Sequence 15282, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15282
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Glycine max
```

```
US-10-953-349-15282

Query Match          50.0%; Score 30; DB 6; Length 77;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      4 VLQNVREVF 12
Db      24 IVQGVRFAP 32

Search completed: June 6, 2006, 00:12:57
Job time : 3.82414 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-27

Perfect score: 61

Sequence: 1 PQGVAQNVRPFV 12

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*
10: geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	12	4	AAB35368 Alpha3bet
2	56	91.8	12	4	AAB35352 Alpha3bet
3	56	91.8	12	4	AAB35378 Alpha3bet
4	56	91.8	12	6	ABG72834 Thrombosop
5	56	91.8	240	8	ADL70641 Human thr
6	56	91.8	432	8	ADQ39359 Human myo
7	56	91.8	432	8	ADQ39357 Human myo
8	56	91.8	459	4	AAU02916 Angiotens
9	56	91.8	466	3	ABAB43602 Human can
10	56	91.8	546	4	AAU02915 Angiotens
11	56	91.8	548	7	ADN02474 TSF polyp
12	56	91.8	555	4	AAU02914 Angiotens
13	56	91.8	731	4	AAU02913 Angiotens
14	56	91.8	1152	3	ABAB0042 Human thr
15	56	91.8	1152	5	ABAB82285 Human thr
16	56	91.8	1152	5	ABAB82285 Human thr
17	56	91.8	1170	4	ABAB74450 Human var
18	56	91.8	1170	4	ABAB90800 Human she
19	56	91.8	1170	5	AAE25030 Human thr
20	56	91.8	1170	5	AAU75315 Human thr
21	56	91.8	1170	6	ABP96780 Human COP
22	56	91.8	1170	6	ABU03474 Angiogene
23	56	91.8	1170	6	ABG74673 Human THB

24	56	91.8	1170	6	AAE36228 Human THB
25	56	91.8	1170	7	ABR62059 Human thr
26	56	91.8	1170	7	ADN39852 Cancer/an
27	56	91.8	1170	8	ADJ76124 Marker ge
28	56	91.8	1170	8	ADJ75296 Marker ge
29	56	91.8	1170	8	ADL70639 Human thr
30	56	91.8	1170	8	ADL35874 Human thr
31	56	91.8	1170	8	ADQ26070 Thrombosop
32	56	91.8	1170	8	ADP54179 Human PRO
33	56	91.8	1170	8	ADQ39358 Human myo
34	56	91.8	1170	8	ADQ39356 Human myo
35	56	91.8	1170	8	ADQ39355 Human myo
36	56	91.8	1170	9	ADZ21688 Thrombosop
37	56	91.8	1170	9	ABE87781 Human thr
38	56	91.8	1170	9	ABE46751 Human thr
39	53	86.9	12	4	AAAB35373 Alpha3bet
40	53	86.9	12	4	AAAB35381 Alpha3bet
41	52	85.2	12	4	AAAB35364 Alpha3bet
42	52	85.2	12	4	AAAB35374 Alpha3bet
43	51	83.6	12	4	AAAB35371 Alpha3bet
44	51	83.6	12	4	AAAB35376 Alpha3bet
45	51	83.6	12	4	AAAB35380 Alpha3bet

ALIGNMENTS

RESULT 1

AAB35368
ID AAB35368 standard; peptide; 12 AA.
XX
AC AAB35368;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3bet1 integrin binding peptide #33.
XX
KW Alpha3bet1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3bet1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
XX of the invention
SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 1 FQGVQNVRFVF 12

RESULT 2
AAB35352
ID AAB35352 standard; peptide; 12 AA.
XX
AC AAB35352;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #17.
XX
KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Example 2; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;
Query Match 91.8%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00034;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 1 FQGVQNVRFVF 12

RESULT 3
AAB35378
ID AAB35378 standard; peptide; 12 AA.
XX
AC AAB35378;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #43.

XX Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Example 2; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;
Query Match 91.8%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00034;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 1 FQGVQNVRFVF 12

RESULT 4
ABG72834
ID ABG72834 standard; peptide; 12 AA.
XX
AC ABG72834;
XX
DT 24-FEB-2003 (first entry)
XX
DE Thrombospondin-1 sequence containing synthetic peptide.
XX
KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
OS Synthetic.
XX
PN WO200281630-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010535.
XX
PR 06-APR-2001; 2001US-0281994P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;

XX WPI; 2003-103329/09.
 XX A new diagnosis for cancer other than prostate cancer in a mammal useful
 PT to detect cancer including lung cancer, particularly small cell lung
 PT cancer and melanoma comprises detecting semenogelin in a sample.
 XX Example 1; Page 14; 32pp; English.
 XX The invention relates to diagnosing cancer other than prostate cancer in
 CC a male mammal, comprising assaying a test sample for increased level of
 CC semenogelin, or cancer in a female by assaying for the presence of
 CC semenogelin. Administering a semenogelin protein or polypeptide fragment
 CC or a semenogelin-specific antibody or active fragment, or a recombinant
 CC vector expressing the protein or antibody, is useful for inducing an
 CC immune response to a cancer in a mammal, where the cancer is not prostate
 CC cancer and semenogelin is a marker. The invention is used to diagnose
 CC cancer, particularly of epithelial origin such as lung cancer, papillary
 CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
 CC (SCLC), or a melanoma. The present sequence represents the amino acid
 CC sequence of the thrombospondin-1 sequence containing synthetic peptide
 XX which binds to alpha-3-beta-1 integrin
 SQ Sequence 12 AA;
 Query Match 91.8%; Score 56; DB 6; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.00034;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVQNVRFVF 12
 DB 1 FQGVQNVRFVF 12
 RESULT 5
 ADL70641
 ID ADL70641 standard; protein; 240 AA.
 AC ADL70641;
 XX
 XX 20-MAY-2004 (first entry)
 XX Human thrombospondin-1 N-terminal domain.
 XX Human; thrombospondin-1; epitope; cancer; diagnosis.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 23..32
 FT /note= "Heparin binding region"
 FT Region 77..82
 FT /note= "Heparin binding region"
 FT Region 151..164
 FT /note= "Fibrinogen binding region"
 XX
 FN WO2004018995-A2.
 XX
 XX 04-MAR-2004.
 XX
 XX 20-AUG-2003; 2003WO-US026023.
 XX
 XX 23-AUG-2002; 2002US-0405494P.
 PR
 XX 21-APR-2003; 2003US-00419462.
 XX
 XX (WILL/) WILLIAMS K J.
 XX
 XX Williams KJ;
 PI
 XX WPI; 2004-226901/21.
 DR
 XX New purified thrombospondin fragment extracted from a body fluid, useful
 PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma

PT or leukemia or as calibrators, indicators, immunogens and analytes.
 XX Disclosure; SEQ ID NO 40; 76pp; English.
 XX The present sequence is that of the N-terminal domain of human
 CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
 CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
 CC clinical assays for cancer and for generation of antibodies and other
 CC binding agents. A method that distinguishes TSP from a TSP fragment or
 CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
 CC or portion as a target for a binding molecule, e.g. an antibody, to
 CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
 CC epitope present in TSP but not in the fragment or portion to obtain a
 CC quantitation of TSP only; and (3) using the difference between (1) and
 CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
 CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
 CC the TSP fragment or portion is performed in order to detect the presence,
 CC or monitor the course, of a disease or condition selected from cancer,
 CC renal failure, renal disease, atopic dermatitis, vasculitis, acute
 CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
 CC infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
 CC nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
 CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
 CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
 CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition
 CC associated with clotting, a condition associated with platelet
 CC activation, a condition associated with intravascular platelet
 CC activation, a condition associated with consumption of platelets, heparin
 CC induced thrombocytopenia, disseminated intravascular coagulation,
 CC intravascular coagulation, extravascular coagulation, a condition
 CC associated with endothelial activation, a condition associated with
 CC production and/or release of thrombospondin and/or a thrombospondin
 CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 CC reaction, an aspartame reaction, atopic dermatitis, eczema,
 CC hypersensitivity, scleroderma, conditions associated with plugging of
 CC vessels, a condition associated with a cryofibrinogen, a condition
 CC associated with a cryoglobulin, and a condition associated with an anti-
 CC cardiolipin antibody. The cancer is selected from adenoma, a
 CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system,
 CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
 CC biliary system, pancreas, head, neck, endocrine system, reproductive
 CC system (male or female), genitourinary system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.
 XX
 SQ Sequence 240 AA;
 Query Match 91.8%; Score 56; DB 8; Length 240;
 Best Local Similarity 91.7%; Pred. No. 0.01;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVQNVRFVF 12
 DB 190 FQGVQNVRFVF 201
 RESULT 6
 ADQ39359
 ID ADQ39359 standard; protein; 432 AA.
 XX
 XX ADQ39359;
 AC
 XX 18-NOV-2004 (first entry)
 DT
 XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.
 DE
 XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 XX cardiac; gene therapy; human.
 KW

```

XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38531.
XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1022; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
XX CC has an altered risk for developing myocardial infarction. The method
XX CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX CC the nucleotide sequences given in the specification in the individual's
XX CC nucleic acids, where the presence of the SNP is correlated with an
XX CC altered risk for myocardial infarction in the individual. The invention
XX CC further comprises: an isolated nucleic acid molecule comprising at least
XX CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
XX CC the specification or its complement and encoding any one of the amino
XX CC acid sequences given in the specification; an isolated polypeptide
XX CC comprising an amino acid sequence given in the specification; an antibody
XX CC that specifically binds to the polypeptide or its antigen-binding
XX CC fragment; an amplified polynucleotide containing an SNP given in the
XX CC specification and which is between about 16 and 1000 nucleotides in
XX CC length; a kit for detecting an SNP in a nucleic acid, comprising the
XX CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
XX CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
XX CC method for identifying an agent useful in treating or preventing
XX CC myocardial infarction. The novel detection method has cardiant activity.
XX CC The nucleic acids of the invention may be used in gene therapy. The
XX CC method is useful in identifying an individual who has an increased or
XX CC decreased risk for developing myocardial infarction and for preparing a
XX CC composition for treating or preventing myocardial infarction. This
XX CC composition represents the protein of a human myocardial infarction-
XX CC associated gene containing one or more SNP's of the invention. Note: This
XX CC sequence was not shown in the specification. The sequence has come from
XX CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;

Query Match          91.8%; Score 56; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

RESULT 7
ID ADQ39357
AC ADQ39357;
XX DT 18-NOV-2004 (first entry)

```

```

XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1020.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human.
XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38529.
XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.
XX PS Claim 10; SEQ ID NO 1020; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
XX CC has an altered risk for developing myocardial infarction. The method
XX CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX CC the nucleotide sequences given in the specification in the individual's
XX CC nucleic acids, where the presence of the SNP is correlated with an
XX CC altered risk for myocardial infarction in the individual. The invention
XX CC further comprises: an isolated nucleic acid molecule comprising at least
XX CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
XX CC the specification or its complement and encoding any one of the amino
XX CC acid sequences given in the specification; an isolated polypeptide
XX CC comprising an amino acid sequence given in the specification; an antibody
XX CC that specifically binds to the polypeptide or its antigen-binding
XX CC fragment; an amplified polynucleotide containing an SNP given in the
XX CC specification and which is between about 16 and 1000 nucleotides in
XX CC length; a kit for detecting an SNP in a nucleic acid, comprising the
XX CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
XX CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
XX CC method for identifying an agent useful in treating or preventing
XX CC myocardial infarction. The novel detection method has cardiant activity.
XX CC The nucleic acids of the invention may be used in gene therapy. The
XX CC method is useful in identifying an individual who has an increased or
XX CC decreased risk for developing myocardial infarction and for preparing a
XX CC composition for treating or preventing myocardial infarction. This
XX CC composition represents the protein of a human myocardial infarction-
XX CC associated gene containing one or more SNP's of the invention. Note: This
XX CC sequence was not shown in the specification. The sequence has come from
XX CC an electronic sequence listing downloaded from the WIPO website.
XX SQ Sequence 432 AA;

Query Match          91.8%; Score 56; DB 8; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

RESULT 8
AAU02916

```



```

RESULT 10
ID AAU02915 standard; protein; 546 AA.
XX
AC AAU02915;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX
OS Homo sapiens.
XX
PN WO200136632-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-IL000766.
XX
PR 17-NOV-1999; 99IL-00132978.
XX
PR 10-DEC-1999; 99IL-00133455.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX
WPI; 2001-336004/35.
XX
DR N-PSDB; AAS06015.
XX
RR Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
XX capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 15; 519pp; English.
XX
CC The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
XX polypeptide receptor 2. The polypeptides and their associated nucleic
XX acids are useful for identification of variant sequences and detection of
XX candidate compounds capable of binding to the molecules. The sequences of
XX the invention can be used in the treatment and diagnosis of various
XX disorders including cardiovascular diseases such as arteriosclerosis,
XX myocardial infarction and coronary arterial thrombosis, renal diseases
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX immune disorders such as immune complex nephritis, multiple sclerosis,
XX cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
XX as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
XX
SQ Sequence 546 AA;
XX
Query Match 91.8%; Score 56; DB 4; Length 546;
Best Local Similarity 91.7%; Pred. No. 0.025;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

```

```

RESULT 11
ID ADN02474 standard; protein; 548 AA.
XX
AC ADN02474;
XX
DT 17-JUN-2004 (first entry)
XX
DE TSF polypeptide.
XX
KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX
OS Homo sapiens.
XX
PN CNI401387-A.
XX
PD 12-MAR-2003.
XX
PF 21-AUG-2002; 2002CN-00129408.
XX
PR 21-AUG-2002; 2002CN-00129408.
XX
PA (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
XX
PI Han Z, Liu P;
XX
WPI; 2003-469302/45.
XX
DR N-PSDB; ADN02475.
XX
PT Tumor suppressing polypeptide TSF and gene therapy vector composition.
XX
PS Claim 2; SEQ ID NO 1; 13pp; Chinese.
XX
CC The present invention relates to a novel recombinant adenovirus vector
XX mediated anti-neoplastic composition is prepared through cloning the CDNA
XX sequence from the human peripheral blood cell by specific primer and
XX reverse transcription-polymerase chain reaction (RT-PCR) method for
XX coding TSF polypeptide, construction in human embryonic kidney cell 293
XX by AdEasy system, and packaging and expressing the recombinant adenovirus
XX vector of TSF. It can suppress the growth and transfer of cancer. The
XX present sequence represents the TSF polypeptide.
XX
SQ Sequence 548 AA;
XX
Query Match 91.8%; Score 56; DB 7; Length 548;
Best Local Similarity 91.7%; Pred. No. 0.025;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

```

```

RESULT 12
ID AAU02914 standard; protein; 555 AA.
XX
AC AAU02914;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

```

KW vascular disorder; asbestosis.

OS Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

PR 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

XX N-PSDB; AAS06014.

XX Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 4; Fig 14; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis

XX Sequence 555 AA;

Query Match 91.8%; Score 56; DB 4; Length 555;

Best Local Similarity 91.7%; Pred. No. 0.026;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRVF 12

DB 208 FQGVQNVRVF 219

RESULT 13

AAU02913

ID AAU02913 standard; protein; 731 AA.

XX AC AAU02913;

XX 12-SEP-2001 (first entry)

XX Angiotensin converting enzyme (ACEV) splice variant protein #13.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

KW vascular disorder; asbestosis.

OS Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

PR 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

XX N-PSDB; AAS06013.

XX Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 4; Fig 13; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis

XX Sequence 731 AA;

Query Match 91.8%; Score 56; DB 4; Length 731;

Best Local Similarity 91.7%; Pred. No. 0.035;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRVF 12

DB 208 FQGVQNVRVF 219

RESULT 14

AA000042

ID AA000042 standard; protein; 1152 AA.

XX AC AA000042;

XX 08-NOV-2000 (first entry)

XX Human thrombospondon-1 (TSP-1).

XX TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
KW diabetic retinopathy; corneal graft rejection; glaucoma.

OS Homo sapiens.

XX Key Location/Qualifiers

XX 361..416

XX Region /label= Type 1 repeat region

```

FT Region 417..473
FT /label= Type 1 repeat region
FT Region 474..530
FT /label= Type 1 repeat region
XX
XX WO200044908-A2.
XX
XX PD 03-AUG-2000.
XX
XX PF 01-FEB-2000; 2000WO-US002482.
XX
XX PR 01-FEB-1999; 99US-0118053P.
XX
XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX PI Lawler JW;
XX
XX DR WPI; 2000-514823/46.
XX
XX PT Nucleic acids encoding chimeric proteins such as cartilage oligomeric
FT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
FT inhibiting angiogenesis and treating diseases such as cancer.
XX
XX PS Disclosure; Fig 1; 40pp; English.
XX
XX CC New nucleic acids are described which encode a protein comprising the
CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
CC the TGF (transforming growth factor)-beta activation region of human TSP-
CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
CC tumours in mice models. Thus the nucleic acids and proteins may be useful
CC for treating angiogenesis related diseases such as cancer (by reducing
CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
CC for treating human immunodeficiency virus (HIV) infection. Anti-
CC angiogenic therapy has little toxicity, does not require the therapeutic
CC agent to enter tumour cells or cross the blood-brain barrier, controls
CC tumour growth independently of growth of tumour cell heterogeneity, and
CC does not induce drug resistance
XX
XX SQ Sequence 1152 AA;
XX
XX Query Match 91.8%; Score 56; DB 3; Length 1152;
XX Best Local Similarity 91.7%; Pred.No. 0.058;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVAQNVRVF 12
XX ||||| |||||
XX Db 190 FQGVQNVRVF 201
XX
XX RESULT 15
XX AAU74771
XX ID AAU74771 standard; protein; 1152 AA.
XX
XX AC AAU74771;
XX
XX DT 09-APR-2002 (first entry)
XX
XX DE Human thrombospondin-1 (TSP-1).
XX
XX KW Thrombospondin-1; TSP-1; cytostatic; angiogenesis; vasotropic; vulnery;
XX KW neovascularisation; cell proliferation inhibitor; cancer; solid tumour;
XX KW haemangioma; acoustic neuromas; neurofibroma; trachoma;
XX KW pyogenic granulomas; rheumatoid arthritis; ocular angiogenic disease;
XX KW retinopathy; psoriasis; macular degeneration; corneal graft rejection;
XX KW neovascular glaucoma; retrolental fibroplasia; rebeosis; angiofibroma;
XX KW Osler-Webber syndrome; myocardial angiogenesis; haemophilic joints;
XX KW plaque neovascularisation; telangiectasia; wound granulation; apoptosis.
XX
XX OS Homo sapiens.

```

```

XX Key Location/Qualifiers
FH Region 263..360
FT /label= Procollagen_homology_region
FT Domain 303..309
FT /label= Procollagen_homology_domain
FT /note= "Required in inhibition of angiogenesis"
FT Domain 361..530
FT /label= Type 1 repeat domain
FT /note= "This region contains 3 type 1 repeats, from
FT residues 361-416, residues 417-473 and residues 474-530"
FT Domain 364..370
FT /label= Heparin_binding_domain
FT Region 413..415
FT /label= RFX motif
FT /note= "Necessary and sufficient for activation of
FT transforming growth factor beta (TGF beta)."
FT Domain 418..423
FT /label= TGF-beta and fibronectin binding_domain
FT /note= "Transforming growth factor"
FT Domain 420..426
FT /label= Heparin_binding_domain
FT Domain 429..434
FT /label= Cell_binding_domain
FT Domain 477..483
FT /label= Heparin_binding_domain
FT Domain 481..499
FT /label= Anti-angiogenesis_domain
FT Domain 486..491
FT /label= Cell_binding_domain
FT Domain 531..673
FT /label= Type 2 repeat domain
FT /note= "This region contains 3 type 2 repeats, from
FT residues 531-571, residues 572-629 and residues 630-673"
FT Domain 570..601
FT /label= Calcium_binding_domain
FT Domain 698..925
FT /label= Type 3 repeat domain
FT /note= "This region contains 7 type 3 repeats, from
FT residues 698-733, residues 734-756, residues 757-792,
FT residues 793-815, residues 816-853, residues 854-889 and
FT residues 890-925"
XX
XX PN WO200191781-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 25-MAY-2001; 2001WO-US017250.
XX
XX PR 26-MAY-2000; 2000US-0207994P.
XX
XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX PI Lawler JW;
XX
XX DR WPI; 2002-106273/14.
XX
XX PT Composition useful for treatment of cancer comprises cDNA encoding amino
FT acids of human thrombospondin-1 or its conservative variant and a
FT carrier.
XX
XX PS Disclosure; Fig 7; 54pp; English.
XX
XX CC The invention describes a composition comprising cDNA encoding fragments
XX of human thrombospondin-1 (TSP-1), a type 1 repeat polypeptide and
XX potent inhibitor of tumour growth and angiogenesis. The composition is
XX useful for killing cancerous cells (preferably tumour); for reducing
XX volume or inhibiting growth of a tumour (inhibiting neovascularisation in
XX the tumour); for decreasing proliferation of tumour cells; in the
XX treatment of diseases and conditions associated with angiogenic activity
XX or misregulated growth and angiogenesis-mediated diseases such as cancer,
XX solid tumour, tumour metastasis, benign tumour, (e.g. haemangioma,
XX acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas),

```

CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases (e.g.
 CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
 CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasias,
 CC rebeosis), Osler-Webber syndrome, myocardial angiogenesis,
 CC telangiectasia, plaque neovascularisation, haemophilic joints,
 CC angiofibroma or wound granulation. The composition induces apoptosis and
 CC inhibits neovascularisation in the tumour cells. This amino acid sequence
 CC represents human thrombospondin-1 (TSP-1), on which the recombinant
 CC proteins of the invention are based

XX
 SQ Sequence 1152 AA;

Query Match 91.8%; Score 56; DB 5; Length 1152;

Best Local Similarity 91.7%; Pred. No. 0.058;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12

Db 190 FQGVQNVRFVF 201

Search completed: June 5, 2006, 22:24:59
 Job time : 92.1379 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-27

Perfect score: 61

Sequence: 1 FQGVQNVRFV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	56	91.8	229	2	S57957	thrombospondin 1 -
2	56	91.8	1170	1	TSHUP1	thrombospondin 1 p
3	56	91.8	1170	2	A40558	thrombospondin 1 p
4	43	70.5	145	2	T20985	hypothetical prote
5	43	70.5	162	2	B88349	protein F15D4_3 (i
6	41	67.2	212	2	D81929	probable imidazole
7	38	62.3	1172	1	TSHUP2	thrombospondin 2 p
8	38	62.3	1172	2	A42587	thrombospondin 2 p
9	37	60.7	332	2	I67791	cytochrome P450 2B
10	36	59.0	295	2	E72462	probable lipoid ac
11	36	59.0	505	2	C69415	carbohydrate kinas
12	36	59.0	586	2	JC5618	beta-galactosidase
13	36	59.0	1178	1	A39804	thrombospondin pre
14	35	57.4	152	2	S44740	C02C2.2 protein -
15	35	57.4	155	2	B83124	hypothetical prote
16	35	57.4	171	2	D95336	hypothetical prote
17	35	57.4	199	1	G64070	imidazoleglycerol
18	35	57.4	308	2	JC5468	leukocidin chain 1
19	35	57.4	311	2	C89968	leukotoxin LuK1 (i
20	35	57.4	444	2	B82057	probable penicilli
21	35	57.4	893	2	T36795	conserved hypothet
22	35	57.4	1207	2	C70013	leukocidin chain F
23	34	55.7	54	2	S35697	hypothetical prote
24	34	55.7	175	2	S75258	gamma-hemolysin co
25	34	55.7	286	2	C49238	synergohymenotropi
26	34	55.7	310	2	S68225	leukocidin chain S
27	34	55.7	312	2	T00160	leukocidin chain S
28	34	55.7	312	2	S32211	leucocidin R S com
29	34	55.7	315	2	A49234	leucocidin R S com

30	34	55.7	315	2	JN0626	leukocidin chain S
31	34	55.7	315	2	E90043	gamma-hemolysin co
32	34	55.7	315	2	PC4078	hlgc-like protein
33	34	55.7	326	2	S28706	hypothetical prote
34	34	55.7	377	1	A26499	regulatory protein
35	34	55.7	531	2	JC5172	probable methylgal
36	34	55.7	548	2	E85860	hypothetical prote
37	34	55.7	548	2	H64990	probable malate de
38	34	55.7	548	2	C91016	probable malate qu
39	34	55.7	556	2	A81931	probable adhesin N
40	34	55.7	652	2	T03504	potassium channel
41	34	55.7	688	2	S55349	probable negative
42	34	55.7	1142	2	T39103	hypothetical prote
43	33	54.1	154	2	T18808	hypothetical prote
44	33	54.1	252	2	F69955	hypothetical prote
45	33	54.1	271	2	S72382	hypothetical prote

ALIGNMENTS

RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C:Accession: S57957

R:Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi

A:Reference number: S57955

A:Accession: S57957

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-229 <LAF>

A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:g899228; PIDN

C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

Query Match 91.8%; Score 56; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0013;

Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 12

|||||

DB 190 FQGVQNVRFV 201

RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)

C>Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A26155; A34274; A30140; A25812; A05172; A42927

C:Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A>Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple ce

A:Reference number: A26155; MUID:87057617; PMID:2430973

A:Accession: A26155

A:Molecule type: mRNA

A:Residues: 1-1170 <LAW>

A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:g37137; PIDN:CI

A>Note: parts of this sequence, including the amino end of the mature protein, were deter

R:Lahterty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A>Title: Characterization of the promoter region of the human thrombospondin gene. DNA se

A:Reference number: A34274; MUID:89291870; PMID:2544587

A:Accession: A34274

A:Molecule type: DNA

A:Residues: 1-166 <LAH>

A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835

R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I.

J. Cell Biol. 108, 729-736, 1989

A>Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

A:Reference number: A30140; MUID:89139590; PMID:2918029
A:Accession: A30140
A:Molecule type: mRNA
A:Residues: 1-83, 'A', '85-522, 'A', '524-1170 <HEN>
A:Cross-references: UNIPARC:UPI0000038AB1; EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PIDN:AAA404
A:Note: parts of this sequence, including the amino end of the mature protein, were determined by analysis of
R:Kobayashi, S.; Eden-McCuthan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A:Title: Partial amino acid sequence of human thrombospondin as determined by analysis of
A:Reference number: A25812; MUID:87157592; PMID:3030396
A:Accession: A25812
A:Molecule type: mRNA
A:Residues: 1-83, 'A', '85-397 <KOB>
A:Cross-references: UNIPARC:UPI000016B0CA; GB:M25631; NID:g538353; PIDN:AAA36741.1; PIDN:AAA404
R:Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A:Reference number: A05172; MUID:86287276; PMID:3461443
A:Accession: A05172
A:Molecule type: mRNA
A:Residues: 1-83, 'A', '85-374, 'RC', <DIX>
A:Cross-references: UNIPARC:UPI000016B140; GB:M14326; NID:g340005; PIDN:AAA61237.1; PIDN:AAA404
A:Note: parts of this sequence, including the amino end of the mature protein, were determined by analysis of
R:Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A:Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A:Reference number: A42927; MUID:92348511; PMID:1379247
A:Accession: A42927
A:Molecule type: protein
A:Residues: 987-1003 <SUN>
A:Cross-references: UNIPARC:UPI00001742C0
A:Note: Cys-992 is shown to have a free sulphydryl
C:Genetics: S68787
A:Gene: GDB:THBS1; TSP1; TSP
A:Cross-references: GDB:120438; OMIM:188060
A:Map position: 15q15-15q15
A:Introns: 23/1
A:Note: the list of introns may be incomplete
C:Complex: homotrimer, disulfide linked
C:Function:
A:Description: participates in cell migration and adhesion, and in platelet aggregation
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology; calcium binding; cell adhesion; glycoprotein; trimer
F:118/Domain: signal sequence #status predicted <SIG>
F:19-1170/Product: thrombospondin 1 #status predicted <MAT>
F:311-375/Domain: von Willebrand factor type C repeat homology <VWC>
F:378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F:551-586/Domain: EGF homology <EGF1>
F:650-689/Domain: EGF homology <EGF2>
F:928-928/Region: cell attachment (R-G-D) motif
F:171-232/Disulfide bonds: #status predicted
F:248,360,708,1067/Binding site: carbohydate (Asn) (covalent) #status predicted
F:270,274/Disulfide bonds: interchain #status predicted
F:610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:1051/Binding site: carbohydate (Asn) (covalent) #status absent

Query Match 91.8%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.0078;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
||| |||||
DB 208 FQGVQNVRFVF 219

RESULT 3
A40558
thrombospondin 1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1992 #sequence revision 05-Jun-1992 #text_change 09-Jul-2004
C:Accession: A40558; A37905; B42587; S68787
R:Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991
A:Title: Characterization of the murine thrombospondin gene.
A:Reference number: A40558; MUID:92128941; PMID:1774063
A:Accession: A40558
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1170 <LAU>
A:Cross-references: UNIPROT:P35441; UNIPARC:UPI0000028012; GB:M62449; GB:M62450; GB:M62462;
1; GB:M62463; GB:M62464; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469
R:Bornstein, P.; Alf, D.; Devarayalu, S.; Framson, P.; Li, P.
J. Biol. Chem. 265, 16691-16698, 1990
A:Title: Characterization of the mouse thrombospondin gene and evaluation of the role of
A:Reference number: A37905; MUID:90375546; PMID:2398070
A:Accession: A37905
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <BOR>
A:Cross-references: UNIPARC:UPI0000160076; GB:J05605; GB:J05606; NID:g201991; PIDN:AAA404
R:Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Selidin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: B42587
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1152, 'P', '1154-1170 <LAH>
A:Cross-references: UNIPARC:UPI0000177A96; GB:M87276
A:Note: sequence extracted from NCBI backbone (NCBIP:81501)
R:Chen, H.; Aeschlimann, D.; Nowien, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A:Title: Expression and initial characterization of recombinant mouse thrombospondin 1 at
A:Reference number: S68787; MUID:96234006; PMID:8654563
A:Accession: S68787
A:Molecule type: protein
A:Residues: 19-26, 'X', '28-37 <CH>
A:Cross-references: UNIPARC:UPI00000177A97
C:Complex: homotrimer, disulfide linked
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology; calcium binding; glycoprotein; homotrimer
F:118/Domain: signal sequence #status predicted <SIG>
F:19-1170/Product: thrombospondin 1 #status predicted <MAT>
F:311-375/Domain: von Willebrand factor type C repeat homology <VWC>
F:378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F:551-586/Domain: EGF homology <EGF>
F:248,360,708,1067/Binding site: carbohydate (Asn) (covalent) #status predicted

Query Match 91.8%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.0078;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
||| |||||
DB 208 FQGVQNVRFVF 219

RESULT 4
T20985
hypothetical protein F15D4.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20985
R:Smyle, R.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19354
A:Accession: T20985
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-145 <WIL>
A:Cross-references: UNIPROT:Q93511; UNIPARC:UPI000007BE33; EMBL:Z80344; PIDN:CA802486.2;
A:Experimental source: clone F15D4
C:Genetics:

A:Gene: CBSP:F15D4.3
A:Map position: 2
A:Introns: 21/3; 82/1

Query Match 70.5%; Score 43; DB 2; Length 145;
Best Local Similarity 58.3%; Pred. No. 0.34;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVVF 12
DB 73 FMGVAQGLRYIF 84

RESULT 5

B88349 protein F15D4.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: B88349

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/

A:Accession: B88349

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-162 <STO>

A:Cross-references: UNIPROT:Q93511; UNIPARC:UPI0000179EEB; GB:chr_II; PIDN:CAB02486.1; F

A:Gene: F15D4.3

A:Map position: 2

Query Match 70.5%; Score 43; DB 2; Length 162;
Best Local Similarity 58.3%; Pred. No. 0.39;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVVF 12
DB 90 FMGVAQGLRYIF 101

RESULT 6

B81929 probable imidazole-glycerol phosphate synthase amidotransferase component (BC 2.4.2.-) N

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C:Accession: D81929

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: D81929

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-212 <PAR>

A:Cross-references: UNIPROT:Q9JVH3; UNIPARC:UPI000012C805; GB:AL162754; GB:AL157959; NID

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: hish; NMA0840

C:Superfamily: amidotransferase hish; trpG homology

C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 67.2%; Score 41; DB 2; Length 212;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVVF 12
DB 144 FQGIQDTRFVF 155

RESULT 7

TSHP2

thrombospondin 2 precursor - human

C:Species: Homo sapiens (man)

C>Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A47379; A42173

R:LaBell, T.L.; Byers, P.H.

Genomics 17, 225-229, 1993

A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: poter

A:Reference number: A47379; MUID:94010892; PMID:8406456

A:Accession: A47379

A:Molecule type: mRNA

A:Residues: 1-1172 <LAB>

A:Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046680; GB:U12350; NID:G307505; PIDN:J

R:LaBell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.

Genomics 12, 421-429, 1992

A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression of

A:Reference number: A42173; MUID:92217961; PMID:1559694

A:Accession: A42173

A:Molecule type: mRNA

A:Residues: 560-1172 <LA2>

A:Cross-references: UNIPARC:UPI00001742C1; GB:M81339

A:Experimental source: fibroblast

A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIIP:95096)

C:Genetics:

A:Gene: GDB:THBS2; TSP2

A:Cross-references: GDB:128789; OMIM:188061

A:Map position: 6q27-6q27

C:Complex: homotrimer, disulfide linked

C:Function:

A:Description: participates in cell migration and adhesion, and in platelet aggregation

C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F:1-16/Domain: signal sequence #status predicted <SIG>

F:19-1172/Product: thrombospondin 2 #status predicted <WAT>

F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>

F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>

F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>

F:493-549/Domain: thrombospondin type 1 repeat homology <THR2>

F:553-588/Domain: EGF homology <EGF1>

F:652-691/Domain: EGF homology <EGF>

F:928-930/Region: cell attachment (R-G-D) motif

F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predict

F:167-226/Disulfide bonds: #status predicted

F:266,270/Disulfide bonds: interchain #status predicted

F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 62.3%; Score 38; DB 1; Length 1172;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVVF 12
DB 202 FRGLQNVLVP 213

RESULT 8

A42587

thrombospondin 2 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A42587; A39851

R:Lahterty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seidlin, M.F.; Dixit, V.M.

J. Biol. Chem. 267, 3274-3281, 1992

A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell

A:Reference number: A42587; MUID:92147683; PMID:1371115

A:Accession: A42587

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-1172 <LAH>

A:Cross-references: UNIPROT:Q03350; UNIPARC:UPI0000029847; GB:L07803; GB:M87275; NID:G341

A:Note: sequence extracted from NCBI backbone (NCBIIP:81502)

R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A:Reference number: A39851; MUID:91302287; PMID:1712771
A:Accession: A39851
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-873 <BOR>
A:CROSS-references: UNIPARC:UPI000016D077; GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C:Keywords: calcium binding; glycoprotein
F:319-377/Domain: von Willebrand factor type C repeat homology <VMC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>

Query Match 62.3%; Score 38; DB 2; Length 1172;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVANVRVVF 12
Db 202 FRGLLNQVHLVF 213
|:|:|:|:|:|

RESULT 9
cytochrome P450 2B15 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I67791
R:Nakayama, K.; Suwa, Y.; Mizukami, Y.; Sogawa, K.; Fujii-Kuriyama, Y.
Gene 136, 333-336, 1993
A:Title: Cloning and sequencing of a novel rat cytochrome P450 2B-encoding gene.
A:Reference number: I53690; MUID:94124025; PMID:8294026
A:Accession: I67791
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-332 <RES>
A:CROSS-references: UNIPROT:Q64583; UNIPARC:UPI0000174D3C; GB:D17349; NID:G468484; PIDN:
C:Genetics:
A:Introns: 53/3; 112/3; 160/1; 222/3; 270/1
C:Superfamily: human cytochrome P450 CYP2B6; cytochrome P450 homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
F:133-296/Domain: cytochrome P450 homology <P45>
F:274/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 60.7%; Score 37; DB 2; Length 332;
Best Local Similarity 54.5%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 QGVANVRVVF 12
Db 277 EGIAQNLFIF 287
|:|:|:|:|:|

RESULT 10
E72462
probable lipic acid synthetase APE2344 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 31-Dec-2004
C:Accession: E72462
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72462
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KAW>

Query Match 59.0%; Score 36; DB 2; Length 295;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVANVRVVF 11
Db 138 FQGVEEHRLV 148
|:|:|:|:|:|

RESULT 11
C69415
carbohydrate kinase, FGGY family homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: C69415
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: C69415
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-505 <KLE>
A:CROSS-references: UNIPROT:O28945; UNIPARC:UPI0000056D86; GB:AE001012; GB:AE000782; NID:
C:Superfamily: xyulokinase

Query Match 59.0%; Score 36; DB 2; Length 505;
Best Local Similarity 54.5%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVANVRVVF 12
Db 382 EGVALNIKWVF 392
|:|:|:|:|:|

RESULT 12
JC5618
beta-galactosidase (EC 3.2.1.23) - Bacillus circulans
N:Alternate names: lactase
C:Species: Bacillus circulans
C:Date: 09-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C:Accession: JC5618
R:Rito, Y.; Sasaki, T.
Biosci. Biotechnol. Biochem. 61, 1270-1276, 1997
A:Title: Cloning and characterization of the gene encoding a novel beta-galactosidase fr
A:Reference number: JC5618; MUID:97446518; PMID:9301106
A:Accession: JC5618
A:Molecule type: DNA
A:Residues: 1-586 <ITO>
A:CROSS-references: UNIPROT:O31341; UNIPARC:UPI00000B3ABF; DDBJ:D88750; NID:g2289789; PII
A:Experimental source: strain ATCC 31382
C:Comment: This enzyme hydrolyzes beta-galactoside bonds.
C:Genetics:
A:Gene: bgac
C:Superfamily: beta-galactosidase bga
C:Keywords: glycosidase; hydrolase

Query Match 59.0%; Score 36; DB 2; Length 586;
Best Local Similarity 54.5%; Pred. No. 41;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVANVRVVF 11
|:|:|:|:|:|

Db 67 FEGIADIVRFI 77

RESULT 13
A39804
thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A39804
R:Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <LAW>
A:Cross-references: UNIPROT:P35440; UNIPARC:UPI00013776D; GB:M60853; NID:G212763; PIDN:
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
F:325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F:386-437/Domain: thrombospondin type 1 repeat homology <THRI>
F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F:499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F:658-697/Domain: EGF homology <EGF>

Query Match 59.0%; Score 36; DB 1; Length 1178;
Best Local Similarity 41.7%; Pred. No. 86;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
|:|:|:|:|
Db 208 FRGLLNQIHLP 219

RESULT 14
S44740
C02C2.2 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S44740
R:Wilson, R.
submitted to the EMBL Data Library, September 1993
A:Description: Sequence of the C. elegans cosmid C02C2.
A:Reference number: S44737
A:Accession: S44740
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <WIL>
A:Cross-references: UNIPROT:P34269; UNIPARC:UPI000013B6F5; EMBL:L23649; NID:G388566; PID
C:Genetics:
A:introns: 33/1
C:Superfamily: Caenorhabditis elegans C02C2.2 protein

Query Match 57.4%; Score 35; DB 2; Length 152;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
|:|:|:|:|
Db 92 YMGVCLNVTFF 103

RESULT 15
B83124
hypothetical protein PA4183 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83124
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-155 <STO>
A:Cross-references: UNIPROT:Q9HWJ8; UNIPARC:UPI00000C5C5C; GB:AE004834; GB:AE004091; NID
C:Genetics:
A:Gene: PA4183

Query Match 57.4%; Score 35; DB 2; Length 155;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QGVQNVRFVF 12
|:|:|:|:|
Db 58 QGVQNVRFVF 68

Search completed: June 5, 2006, 22:45:01
Job time : 14.9655 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 108.931 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-27

Perfect score: 61

Sequence: 1 FQGVQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	56	91.8	229	2	Q28194_BOVIN	Q28194 bos taurus
2	56	91.8	496	2	Q7SY84_XENLA	Q7SY84 xenopus lae
3	56	91.8	1170	1	TSP1_BOVIN	Q28178 bos taurus
4	56	91.8	1170	1	TSP1_HUMAN	P07996 homo sapien
5	56	91.8	1170	1	TSP1_MOUSE	P35441 mus musculus
6	56	91.8	1170	2	Q3TR40_MOUSE	Q3tr40 mus musculus
7	56	91.8	1170	2	Q71SA3_RAT	Q71sa3 ratus norv
8	56	91.8	1171	2	Q80YQ1_MOUSE	Q80yq1 mus musculus
9	56	91.8	1171	2	Q8CGB2_MOUSE	Q8cgb2 mus musculus
10	56	91.8	1173	1	TSP1_XENLA	P35448 xenopus lae
11	56	91.8	1225	2	Q59E99_HUMAN	Q59e99 homo sapien
12	51	83.6	1090	2	Q5SPG5_BRARE	Q5spg5 brachydanio
13	51	83.6	1193	2	Q4S758_TETNG	Q4s758 tetraodon n
14	50	82.0	249	2	Q5U903_PIG	Q5u903 sus scrofa
15	50	82.0	1171	2	Q4RLR5_TETNG	Q4rlr5 tetraodon n
16	46	75.4	212	2	Q5FA21_NEIG1	Q5fa21 neisseria g
17	46	75.4	1034	2	Q4RQ74_TETNG	Q4rq74 tetraodon n
18	41	67.2	212	1	H1SS_NEIMA	Q9jvh3 neisseria m
19	41	67.2	727	2	Q8PRY3_METNA	Q8pry3 methanosarc
20	41	67.2	730	2	Q467V1_METBA	Q467v1 methanosarc
21	41	67.2	733	2	Q8TLX6_METAC	Q8tlx6 methanosarc
22	40	65.6	394	2	Q87M68_VIBPA	Q87m68 vibrio para
23	40	65.6	606	2	Q36CS3_9GAMM	Q36cs3 shewanella
24	40	65.6	610	2	Q22B87_9GAMM	Q22b87 shewanella
25	40	65.6	610	2	Q35YNG_9GAMM	Q35yn6 shewanella
26	40	65.6	619	2	Q2X146_9GAMM	Q2x146 shewanella
27	40	65.6	619	2	Q3QA31_9GAMM	Q3qa31 shewanella
28	40	65.6	621	2	Q7RGF7_PLAYO	Q7rgf7 plasmodium
29	40	65.6	629	2	Q2ZSB6_SHEPU	Q2zsb6 shewanella
30	39	63.9	289	2	Q48AF4_COLP3	Q48af4 colwellia p
31	39	63.9	344	2	Q5AWH7_EMENTI	Q5awh7 aspergillus

32	39	63.9	411	2	Q3XC93_METFL	Q3xc93 methylobaci
33	39	63.9	500	2	Q33FLO_METHU	Q33f10 methanospir
34	39	63.9	859	2	Q6BY75_DEBHA	Q6by75 debaryomyce
35	39	63.9	1113	2	Q2LTC9_9DELT	Q2ltc9 syntrophus
36	39	63.9	1168	2	Q5VH52_CIOIN	Q5vh52 ciona intes
37	39	63.9	1720	2	Q314W5_DESDG	Q314w5 desulfovibr
38	38	62.3	29	2	Q4XGT3_PLACH	Q4xgt3 plasmodium
39	38	62.3	253	2	Q4IUL7_AZOVI	Q4iul7 azorobacter
40	38	62.3	261	2	Q6D444_ERWCT	Q6d444 erwania car
41	38	62.3	508	2	Q6PFY9_9PROT	Q6ppy9 magnetite-c
42	38	62.3	1172	1	TSP2_HUMAN	P35442 homo sapien
43	38	62.3	1172	1	TSP2_MOUSE	Q03350 mus musculu
44	38	62.3	1172	2	Q5RI52_HUMAN	Q5ri52 homo sapien
45	38	62.3	1172	2	Q7TMT3_MOUSE	Q7tmt3 mus musculu

ALIGNMENTS

RESULT 1
Q28194_BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=963331130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCP19>3.3.CO;2-0;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
RT J. Cell. Physiol. 167:164-172(1996).
RL -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON_TER 1
FT NON_TER 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCAE4B6B69C CRC64;

Query Match 91.8%; Score 56; DB 2; Length 229;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
||| |||||
Db 190 FQGVQNVRFVF 201

RESULT 2
Q7SY84_XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
ID Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.

DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBSJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC054970; AAH54970.1; -; mRNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR001007; WVF_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; WVC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; WVF_C; 1.
DR PROSITE; PS0184; WVF_C; 1.
SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7EF51B CRC64;

Query Match 91.8%; Score 56; DB 2; Length 496;
Best Local Similarity 91.7%; Pred. No. 0.025; 1; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 1;
1 FQGVQNVRFVF 12
|||||

Db 214 FQGVQNVRFVF 225
RESULT 3
TSP1_BOVIN
ID TSP1_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=Holstein; TISSUE=Tooth;
RX MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
RT odontoblasts and predentin.";
RL Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis
CC and/or maintenance of dentin and dental pulp.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- TISSUE SPECIFICITY: Odontoblasts.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 WVF domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB005287; BAA21115.1; -; mRNA.
DR EMBL; X87618; CAA60950.1; -; mRNA.
DR EMBL; X87619; CAA60951.1; -; mRNA.
DR PIR; S5501; S55501.
DR HSP; P07896; 1LSL.
DR SMR; Q28178; 549-1169.
DR GlycoSuiteDB; Q28178; --
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF 3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR013032; EGF-like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; WVF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 13.

DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS0026; EGF_3; 2.
DR PROSITE; PS0092; TSP1; 3.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS0184; VWC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 646
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT REGION 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT CARBOHYD 1085 1085
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT DISULFID 519 531
FT DISULFID 551 562
FT DISULFID 556 572
FT DISULFID 575 586
FT DISULFID 592 608
FT DISULFID 599 617
FT DISULFID 620 644
FT DISULFID 650 663
FT DISULFID 657 676
FT DISULFID 678 689
FT DISULFID 705 713
FT DISULFID 718 738
FT DISULFID 754 774
FT DISULFID 777 797
FT DISULFID 813 833
FT DISULFID 836 856
FT DISULFID 874 894
FT DISULFID 910 930
FT DISULFID 946 1167
FT CONFLICT 805 805 S -> G (in Ref. 2).
SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 91.8%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred.No. 0.061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVQNVRFVF 12
DB 208 FQGVQNVRFVF 219
RESULT 4
TSP1_HUMAN
ID TSP1_HUMAN STANDARD; PRT; 1170 AA.
AC P07596; Q15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DT 07-MAR-2006, entry version 78.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologues with several different
RT proteins.";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier W.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region.";
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Franson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins.";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;

RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Ratalinic J.,
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT module.";
RL J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RT of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hyrazide chemistry, and mass spectrometry.";
RL J. Proteome Res. 4:2070-2080(2005).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3.
CC -!- SURUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04655; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAA61237.1; ALT SEQ; mRNA.
DR EMBL; J04935; AAA61178.1; -; Genomic_DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1ZA4; X-ray; A=19-257.
DR PDB; 2BRF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000894; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; TSP_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWF_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01184; VWFC_2; 1.
KW 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTid=PRO_0000035842.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal. (Potential).
FT Heparin-binding site (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...).
FT N-linked (GlcNAc...). (Potential).
FT C-linked (Man).
FT /FTid=CAR_000205.
FT O-linked (Fuc...).
FT /FTid=CAR_000206.
FT C-linked (Man).
FT /FTid=CAR_000207.
FT C-linked (Man).
FT /FTid=CAR_000208.
FT O-linked (Fuc...).
FT /FTid=CAR_000209.
FT C-linked (Man).
FT /FTid=CAR_000210.
FT O-linked (Fuc...).
FT /FTid=CAR_000211.
FT N-linked (GlcNAc...). (Potential).
FT Interchain (Probable).
FT Interchain (Probable).

Query Match 91.8%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%;
Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 FQGVQNVRFVF 12
DB 208 FQGVQNVRFVF 219

RESULT 5

TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 57.
DE Thrombospondin-1 precursor.
GN Name=Thbs1; Synonyms=Tspl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92147883; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin 2 sequence and expression during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92147883; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-490.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfai D., Devarayalu S., Framson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
RN [4]
RP PROTEIN SEQUENCE OF 19-37.
RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
RT "Expression and initial characterization of recombinant mouse thrombospondin 1 and thrombospondin 3.";
RL FEBS Lett. 387:36-41(1996).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibrinogen, fibronectin, laminin, type V collagen and integrins alpha-V/beta-1, alpha-V/beta-3 and alpha-IIB/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; M62470; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62450; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62451; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62452; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62453; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62454; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62455; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62456; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62457; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62458; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62459; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62460; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62461; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62462; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62463; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62464; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62465; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62466; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62467; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62468; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M62469; AA50611.1; JOINED; Genomic DNA.
DR EMBL; M87276; AA53063.1; -; mRNA.
DR EMBL; J05606; AAA40431.1; -; Genomic DNA.
DR PIR; A40558; A40558.
DR HSRP; P07996; LLSL.
DR SMR; P35441; 549-1169.
DR Ensemble; ENSMUSG00000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF 3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR013032; EGF-like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP 1; 3.
DR Pfam; PF02412; TSP 3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF 1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC 1; 1.
DR PROSITE; PS0184; VWFC 2; 1.
KW Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
KW Glycoprotein; Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO 0000035843.
FT TSP N-terminal.
FT VWFC
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708

```

FT CARBOHYD 1067 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 Interchain (Probable).
FT DISULFID 274 Interchain (Probable).
FT DISULFID 274 By similarity.
FT DISULFID 391 By similarity.
FT DISULFID 395 By similarity.
FT DISULFID 406 By similarity.
FT DISULFID 447 By similarity.
FT DISULFID 451 By similarity.
FT DISULFID 462 By similarity.
FT DISULFID 474 By similarity.
FT DISULFID 504 By similarity.
FT DISULFID 508 By similarity.
FT DISULFID 519 By similarity.
FT DISULFID 551 By similarity.
FT DISULFID 556 By similarity.
FT DISULFID 572 By similarity.
FT DISULFID 586 By similarity.
FT DISULFID 592 By similarity.
FT DISULFID 599 By similarity.
FT DISULFID 617 By similarity.
FT DISULFID 620 By similarity.
FT DISULFID 650 By similarity.
FT DISULFID 663 By similarity.
FT DISULFID 657 By similarity.
FT DISULFID 678 By similarity.
FT DISULFID 705 By similarity.
FT DISULFID 718 By similarity.
FT DISULFID 738 By similarity.
FT DISULFID 754 By similarity.
FT DISULFID 774 By similarity.
FT DISULFID 813 By similarity.
FT DISULFID 833 By similarity.
FT DISULFID 856 By similarity.
FT DISULFID 874 By similarity.
FT DISULFID 910 By similarity.
FT DISULFID 946 By similarity.
FT DISULFID 961 By similarity.
FT CONFLICT 1025 F -> L (in Ref. 2).
SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 91.8%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRVVF 12
DB 208 FQGVQNVRVVF 219

RESULT 6
Q3TR40 MOUSE PRELIMINARY; PRT; 1170 AA.
AC Q3TR40;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-PEB-2006, entry version 5.
DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DE clone:A530055N06 product:thrombospondin 1, full insert sequence.
GN Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impiombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,

```

```

RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Heminger P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnietz L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone P., Miyake S., Morris K.,
RA Mottaqui-Tabar S., Mulder N., Nakano N., Nakouchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wuestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Pukada S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome."
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schraml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yamanishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```

```

RT 60,770 full-length cDNAs." ;
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Moshina J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes." ;
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer." ;
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK163092; BAE37190.1; -; mRNA.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR006210; EGF 3.
DR InterPro; IPR00742; EGF.
DR InterPro; IPR01881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003167; tsp_3.
DR InterPro; IPR008859; TSP_C.

```

```

DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF 2; UNKNOWN_1.
DR PROSITE; PS50026; EGF 3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C; 1.
SQ SEQUENCE 1170 AA; 129619 MW; 55BF04A20D91B194 CRC64;

Query Match          91.8%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVF 12
Db 208 FOGVLQNVRFV 219
|||||
|||||

RESULT 7
Q71SA3 RAT Q71SA3 RAT PRELIMINARY; PRT; 1170 AA.
AC Q71SA3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Thrombospondin 1.
GN Name=Tsp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RA Iwabuchi A., Hirohata S., Kusachi S., Nakamura K., Murakami T.,
RA Ninomiya Y., Tsuji T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF309630; AAQ14549.1; -; mRNA.
DR SMR; Q71SA3; 834-1169.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
DR InterPro; IPR013320; Cona_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF 3.
DR InterPro; IPR01881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003167; tsp_3.
DR InterPro; IPR008859; TSP_C.

```


RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor
RC arose spontaneously;
RC genome sequenced;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels B., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Heminger P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humenick L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Knapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Read J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
RA Tannoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki K., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kohi M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RA RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayateu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino Y.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630(2000).
RN [9]

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary Gland;
RA MEDLINE=2053013; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[10]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary Gland;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; BC042422; AAH42422.1; -; mRNA.
DR EMBL; AK145202; BAE26293.1; -; mRNA.
DR HSSP; P07996; 1LSL.

Query Match 91.8%; Score 56; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
    |||||
DB 208 FQGVQNVRFVF 219

RESULT 10
TSPI_XENLA
ID TSPI_XENLA STANDARD; PRT; 1173 AA.
AC P35448.
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 54.
DE Thrombospondin-1 precursor.
GN Name=thbs1; Synonyms=tspl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RT "Cloning, characterization and expression of thrombospondin-1 in
RT Xenopus laevis embryos.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-1/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domains.
CC -!- SIMILARITY: Contains 3 TSP type-3 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC

```

```

CC CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
DR HSSP; P07996; 1LSL.
DR SMR; P35448; 552-1172.
DR InterPro; IPR013320; Cona_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_Tsp_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; TSP_3.
DR InterPro; IPR008859; TSP_C.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1RPEPT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1173 Thrombospondin-1.
FT DOMAIN 23 224 /FTid=PRO_0000035844.
FT DOMAIN 319 376 TSP N-terminal.
FT DOMAIN 382 432 TSP type-1 1.
FT DOMAIN 438 493 TSP type-1 2.
FT DOMAIN 495 550 TSP type-1 3.
FT DOMAIN 550 590 EGF-like 1.
FT DOMAIN 591 648 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 649 693 EGF-like 3.
FT DOMAIN 726 761 TSP type-3 1.
FT DOMAIN 762 784 TSP type-3 2.
FT DOMAIN 785 820 TSP type-3 3.
FT DOMAIN 821 843 TSP type-3 4.
FT DOMAIN 844 881 TSP type-3 5.
FT DOMAIN 882 917 TSP type-3 6.
FT DOMAIN 918 953 TSP type-3 7.
FT DOMAIN 954 1173 TSP C-terminal.
FT REGION 23 235 Heparin-binding (Potential).
FT MOTIF 929 931 Cell attachment site (Potential).
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 363 363 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 705 705 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 711 711 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1070 1070 N-linked (GlcNAc...) (Potential).
FT DISULFID 394 426 By similarity.
FT DISULFID 398 431 By similarity.
FT DISULFID 409 416 By similarity.
FT DISULFID 450 487 By similarity.
FT DISULFID 454 492 By similarity.
FT DISULFID 465 477 By similarity.
FT DISULFID 507 544 By similarity.
FT DISULFID 511 549 By similarity.
FT DISULFID 522 534 By similarity.

```

```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 602 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 91.8%; Score 56; DB 1; Length 1173;
Best Local Similarity 91.7%; Pred. No. 0.061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
Db 211 FQGVQNVRFVF 222

RESULT 11
Q59E99 HUMAN
ID Q59E99_HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RA "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB209912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939, 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
```

```
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9889B16E57157B12 CRC64;

Query Match 91.8%; Score 56; DB 2; Length 1225;
Best Local Similarity 91.7%; Pred. No. 0.064;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
Db 263 FQGVQNVRFVF 274

RESULT 12
Q5SPG5 BRARE
ID Q5SPG5_BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (by similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AL928866; CAI20599.1; -; Genomic_DNA.
DR SMR; Q5SPG5; 751-804, 754-1089.
DR Ensembl; ENSDARG0000010785; Danio rerio.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
```

```
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 2.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;

Query Match 83.6%; Score 51; DB 2; Length 1090;
Best Local Similarity 83.3%; Pred. No. 0.6;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQVQNRVVF 12
Db 185 FMGVQLQNRVVF 196

RESULT 13
Q4S758.TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, entry version 1.
DE Chromosome 14 SCAP14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00022976001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99983;
RN NUCLEOTIDE SEQUENCE.
RX PubMed=1549614; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouteau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Bismont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.-P., Guzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
-----
```

```
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; CAAE01014723; CAG03524.1; -; Genomic_DNA.
DR SMR; Q4S758; 811-1148.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS50184; VWF_C; 1.
KW Cell adhesion.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 83.6%; Score 51; DB 2; Length 1193;
Best Local Similarity 83.3%; Pred. No. 0.66;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQVQNRVVF 12
Db 187 FMGVQLQNRVVF 198

RESULT 14
Q5U903.PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN NUCLEOTIDE SEQUENCE.
RP Zhang K., Mauc G., Hauet T.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY773342; AAV38110.1; -; mRNA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWF_C.
-----
```



```
DR Pfam: PF00090; TSP_1; 2.
DR Pfam: PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS0184; VWC_2; 1.
FT NON_TER 1
FT NON_TER 249
SQ SEQUENCE 249 AA; 27560 MW; 465D664BE0329C32 CRC64;

Query Match 82.0%; Score 50; DB 2; Length 249;
Best Local Similarity 90.9%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQNVRFVF 12
Db 1 QGVQNVRFVF 11

RESULT 15
Q4RLR5_TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Chromosome 10 SCAR15019, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00032374001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
CC EMBL; CAAE01015019; CNG10667.1; -; Genomic_DNA.
DR SMR; Q4RLR5; 834-887, 837-1171.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

```
DR InterPro: IPR006210; EGF.
DR InterPro: IPR000742; EGF_3.
DR InterPro: IPR001881; EGF_Ca_bd.
DR InterPro: IPR013032; EGF_like_reg.
DR InterPro: IPR003129; Laminin_G_TSP_N.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR008085; TSP 1.
DR InterPro: IPR001007; VWC_C.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00266; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWC_1; 1.
DR PROSITE; PS0184; VWC_2; 1.
KW Cell adhesion.
FT NON_TER 1171
FT NON_TER 1171
SQ SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 82.0%; Score 50; DB 2; Length 1171;
Best Local Similarity 83.3%; Pred. No. 1;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FGVQNVRFVF 12
Db 207 FTGVLQNVRFVF 218

Search completed: June 5, 2006, 22:42:41
Job time : 109.931 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-27

Perfect score: 61

Sequence: 1 FQGVQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5 COMB.pdp.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6 COMB.pdp.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7 COMB.pdp.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H COMB.pdp.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PTUS COMB.pdp.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE COMB.pdp.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	91.8	825	2	US-09-949-002-482
2	56	91.8	831	2	US-09-939-853A-97
3	56	91.8	831	2	US-09-939-853A-98
4	56	91.8	1170	1	US-08-313-288B-20
5	56	91.8	1170	2	US-09-657-472-2
6	56	91.8	1170	2	US-09-949-002-350
7	38	62.3	1045	2	US-09-949-016-11112
8	38	62.3	1172	1	US-08-313-288B-19
9	38	62.3	1172	2	US-09-949-016-6333
10	35	57.4	206	2	US-09-252-991A-20515
11	35	57.4	2475	2	US-09-413-814-48
12	34	55.7	61	2	US-09-270-767-61396
13	34	55.7	227	2	US-09-303-518D-460
14	34	55.7	304	2	US-09-270-767-45864
15	34	55.7	347	2	US-09-248-796A-19467
16	34	55.7	424	1	US-08-419-414-2
17	34	55.7	436	2	US-09-252-991A-28888
18	34	55.7	440	1	US-08-307-499-15
19	34	55.7	440	2	US-09-299-268-15
20	34	55.7	529	2	US-09-801-042-2
21	34	55.7	556	2	US-09-303-518D-466
22	34	55.7	628	2	US-09-303-518D-464
23	34	55.7	684	3	US-10-124-807-19
24	34	55.7	684	3	US-10-124-807-24
25	34	55.7	731	2	US-09-270-767-42057
26	33	54.1	40	1	US-07-868-353A-3

27	33	54.1	40	1	US-08-407-804-3	Sequence 3, Appl1
28	33	54.1	40	2	US-09-124-807-3	Sequence 3, Appl1
29	33	54.1	171	2	US-09-328-352-8227	Sequence 8227, Ap
30	33	54.1	213	2	US-09-489-039A-14310	Sequence 14310, A
31	33	54.1	327	2	US-09-134-000C-6682	Sequence 6682, Ap
32	33	54.1	345	2	US-09-543-681A-4256	Sequence 4256, Ap
33	33	54.1	345	2	US-09-134-000C-6576	Sequence 6576, Ap
34	33	54.1	350	1	US-07-868-353A-14	Sequence 14, Appl
35	33	54.1	350	1	US-08-407-804-23	Sequence 23, Appl
36	33	54.1	350	2	US-09-124-807-23	Sequence 23, Appl
37	33	54.1	354	1	US-07-868-353A-12	Sequence 12, Appl
38	33	54.1	354	1	US-07-868-353A-13	Sequence 13, Appl
39	33	54.1	354	1	US-07-868-353A-15	Sequence 15, Appl
40	33	54.1	354	1	US-08-407-804-21	Sequence 21, Appl
41	33	54.1	354	1	US-08-407-804-22	Sequence 22, Appl
42	33	54.1	354	1	US-08-407-804-24	Sequence 24, Appl
43	33	54.1	354	2	US-09-124-807-21	Sequence 21, Appl
44	33	54.1	354	2	US-09-124-807-22	Sequence 22, Appl
45	33	54.1	354	2	US-09-124-807-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-482

Query Match 91.8%; Score 56; DB 2; Length 825;

Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12

DB 308 FQGVQNVRFVF 319

RESULT 2

US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.8%; Score 56; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQNVRFVF 12
      ||||| |||||
Db      208 FQGVQNVRFVF 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          91.8%; Score 56; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQNVRFVF 12
      ||||| |||||
Db      208 FQGVQNVRFVF 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          91.8%; Score 56; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQNVRFVF 12
      ||||| |||||
Db      208 FQGVQNVRFVF 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          91.8%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVQNVRFVF 12
      ||||| |||||
Db      208 FQGVQNVRFVF 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-350

Query Match          91.8%; Score 56; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
   |||| |||||
Db 208 FQGVQNVRFVF 219

RESULT 7
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11112

Query Match          62.3%; Score 38; DB 2; Length 1045;
Best Local Similarity 58.3%; Pred. No. 57;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
   ||: ||| ||
Db 281 FRGLQNVHLVF 292

RESULT 8
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
```

```
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match          62.3%; Score 38; DB 1; Length 1172;
Best Local Similarity 58.3%; Pred. No. 65;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
   ||: ||| ||
Db 202 FRGLQNVHLVF 213

RESULT 9
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6333

Query Match          62.3%; Score 38; DB 2; Length 1172;
Best Local Similarity 58.3%; Pred. No. 65;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFVF 12
   ||: ||| ||
Db 202 FRGLQNVHLVF 213

RESULT 10
```

US-09-252-991A-20515
; Sequence 20515, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252.991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 20515

; LENGTH: 206

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20515

Query Match 57.4%; Score 35; DB 2; Length 206;

Best Local Similarity 63.6%; Pred. No. 36;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVAQNVRFVF 12

||||:||||

Db 109 QGVAHDMRFDF 119

RESULT 11

US-09-413-814-48

; Sequence 48, Application US/09413814

; Patent No. 6225064

; GENERAL INFORMATION:

; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH

; APPLICANT: Bristol-Myers Squibb, Co.

; APPLICANT: Beyer, Stefan

; APPLICANT: Bloecker, Helmut

; APPLICANT: Brandt, Petra

; APPLICANT: Cino, Paul M

; APPLICANT: Dougherty, Brian A

; APPLICANT: Goldberg, Steven L

; APPLICANT: Hofle, Gerhard

; APPLICANT: Mueller, Joachim

; APPLICANT: Reichenbach, Hans

; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or

; TITLE OF INVENTION: heteropolypeptide compounds

; FILE REFERENCE: PCV/US 99/23535

; CURRENT APPLICATION NUMBER: US/09/413,814

; CURRENT FILING DATE: 1999-10-07

; EARLIER APPLICATION NUMBER: DE 198 46 493.2

; EARLIER FILING DATE: 1998-10-09

; NUMBER OF SEQ ID NOS: 107

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 48

; LENGTH: 2475

; TYPE: PRT

; ORGANISM: Sorangium cellulosum

US-09-413-814-48

Query Match 57.4%; Score 35; DB 2; Length 2475;

Best Local Similarity 50.0%; Pred. No. 6.1e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGAQNVRFVF 12

||||:||||

Db 123 FAGVSSNLSFLF 134

RESULT 12

US-09-270-767-61396

; Sequence 61396, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 61396

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-270-767-61396

Query Match 55.7%; Score 34; DB 2; Length 61;

Best Local Similarity 58.3%; Pred. No. 14;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGAQNVRFVF 12

||||:||||

Db 44 FPGVTVNRLFP 55

RESULT 13

US-09-303-518D-460

; Sequence 460, Application US/09303518D

; Patent No. 691431

; GENERAL INFORMATION:

; APPLICANT: Scarlato, Vincenzo

; APPLICANT: Masignani, Vega

; APPLICANT: Rappuoli, Rino

; APPLICANT: Pizza, Mariagrazia

; APPLICANT: Grandi, Guido

; TITLE OF INVENTION: Neisserial Antigens

; FILE REFERENCE: CHIR0160

; CURRENT APPLICATION NUMBER: US/09/303,518D

; CURRENT FILING DATE: 1999-04-30

; NUMBER OF SEQ ID NOS: 1098

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 460

; LENGTH: 227

; TYPE: PRT

; ORGANISM: Neisseria meningitidis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (12)..(12)

; OTHER INFORMATION: Xaa= any amino acid

US-09-303-518D-460

Query Match 55.7%; Score 34; DB 2; Length 227;

Best Local Similarity 75.0%; Pred. No. 64;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVAQNV 9

:|||||:

Db 151 KGVAQNI 158

RESULT 14

US-09-270-767-45864

; Sequence 45864, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45864

; LENGTH: 304

```
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45864

Query Match          55.7%; Score 34; DB 2; Length 304;
Best Local Similarity 58.3%; Pred. No. 89;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 FQVQNVRVF 12
Db      287 FPGVTNRLLF 298

RESULT 15
US-09-248-796A-19467
; Sequence 19467, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19467
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19467

Query Match          55.7%; Score 34; DB 2; Length 347;
Best Local Similarity 77.8%; Pred. No. 16+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 GVAQNVRVF 11
Db      217 GVAKNVRIV 225

Search completed: June 5, 2006, 22:48:54
Job time : 23.8966 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-27

Perfect score: 61

Sequence: 1 FQGVQNVRFVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	56	91.8	12	4	US-10-474-213-28
2	56	91.8	240	4	Sequence 28, Appl
3	56	91.8	240	5	Sequence 40, Appl
4	56	91.8	432	5	Sequence 1020, Ap
5	56	91.8	432	5	Sequence 1022, Ap
6	56	91.8	459	6	Sequence 1047, Ap
7	56	91.8	466	3	Sequence 454, App
8	56	91.8	555	6	Sequence 456, App
9	56	91.8	578	6	Sequence 452, App
10	56	91.8	685	6	Sequence 453, App
11	56	91.8	804	6	Sequence 455, App
12	56	91.8	828	6	Sequence 97, Appl
13	56	91.8	831	3	Sequence 461, App
14	56	91.8	831	3	Sequence 457, App
15	56	91.8	855	6	Sequence 458, App
16	56	91.8	1000	6	Sequence 1, Appli
17	56	91.8	1105	6	Sequence 7, Appli
18	56	91.8	1150	3	Sequence 12, Appl
19	56	91.8	1152	3	Sequence 2, Appli
20	56	91.8	1169	5	Sequence 114, App
21	56	91.8	1170	4	Sequence 1170, Ap
22	56	91.8	1170	4	Sequence 38, Appl
23	56	91.8	1170	4	Sequence 482, App
24	56	91.8	1170	4	Sequence 482, App
25	56	91.8	1170	4	Sequence 482, App
26	56	91.8	1170	4	Sequence 482, App
27	56	91.8	1170	4	Sequence 482, App

28	56	91.8	1170	4	US-10-419-462-38	Sequence 38, Appl
29	56	91.8	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	56	91.8	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	56	91.8	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	56	91.8	1170	5	US-10-782-968-38	Sequence 38, Appl
33	56	91.8	1170	5	US-10-849-989-44	Sequence 44, Appl
34	56	91.8	1170	5	US-10-631-467-548	Sequence 548, App
35	56	91.8	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	56	91.8	1170	5	US-10-831-997-2	Sequence 2, Appli
37	56	91.8	1170	5	US-10-995-561-594	Sequence 594, App
38	56	91.8	1170	5	US-10-995-561-595	Sequence 595, App
39	56	91.8	1170	5	US-10-995-561-596	Sequence 596, App
40	56	91.8	1170	6	US-11-037-713-51	Sequence 51, Appl
41	56	91.8	1170	6	US-11-046-644-28	Sequence 28, Appl
42	56	91.8	1170	6	US-11-046-456-28	Sequence 28, Appl
43	46	75.4	226	5	US-10-467-657-2428	Sequence 153, App
44	40	65.6	15	4	US-10-285-394-153	Sequence 6, Appli
45	38	62.3	16	3	US-09-822-682-6	

ALIGNMENTS

RESULT 1

US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Kruttsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/10535
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 91.8%; Score 56; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00033;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12

Db 1 FQGVQNVRFVF 12

RESULT 2

US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof in Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 40

; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match 91.8%; Score 56; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.0099;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 190 FQGVQNVRFVF 201

RESULT 3

US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match 91.8%; Score 56; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.0099;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 190 FQGVQNVRFVF 201

RESULT 4

US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 91.8%; Score 56; DB 5; Length 432;

Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQNVRFVF 219

RESULT 5

US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 91.8%; Score 56; DB 5; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQNVRFVF 219

RESULT 6

US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: Thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match 91.8%; Score 56; DB 6; Length 459;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
| | | | | | | | | |
Db 208 FQGVQNVRFVF 219

RESULT 7

US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 91.8%; Score 56; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVF 12
||| |||||
Db 261 FOGVLQNVRFV 272

RESULT 8

US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:

; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 91.8%; Score 56; DB 6; Length 555;
Best Local Similarity 91.7%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVF 12
||| |||||
Db 208 FOGVLQNVRFV 219

RESULT 9

US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:

; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 91.8%; Score 56; DB 6; Length 578;
Best Local Similarity 91.7%; Pred. No. 0.027;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVF 12
||| |||||
Db 208 FOGVLQNVRFV 219

RESULT 10

US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:

; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 91.8%; Score 56; DB 6; Length 685;
Best Local Similarity 91.7%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVF 12
||| |||||
Db 208 FOGVLQNVRFV 219

RESULT 11

US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:

; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 91.8%; Score 56; DB 6; Length 804;
Best Local Similarity 91.7%; Pred. No. 0.039;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FOGVAQNVRVF 12
||| |||||
Db 208 FOGVLQNVRFV 219

RESULT 12

US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:

; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match          91.8%; Score 56; DB 6; Length 828;
Best Local Similarity 91.7%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.8%; Score 56; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-09-939-853A-98

Query Match          91.8%; Score 56; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.041;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match          91.8%; Score 56; DB 6; Length 855;
Best Local Similarity 91.7%; Pred. No. 0.042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVQNVRFVF 12
Db 208 FQGVQNVRFVF 219

Search completed: June 6, 2006, 00:00:10
Job time : 78.6207 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-27

Perfect score: 61

Sequence: 1 FQVAGNVRVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB_PEP.*
- 2: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB_PEP.*
- 3: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB_PEP.*
- 4: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB_PEP.*
- 5: /EMC Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB_PEP.*
- 6: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB_PEP.*
- 7: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB_PEP.*
- 8: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	54.1	608	7	US-11-203-828-5
2	32.5	53.3	157	6	US-10-953-349-6894
3	32.5	53.3	208	6	US-10-953-349-6893
4	32.5	53.3	263	6	US-10-953-349-6892
5	32	52.5	980	7	US-11-242-505A-36
6	31	50.8	164	6	US-10-953-349-1159
7	31	50.8	249	6	US-10-953-349-1158
8	31	50.8	250	6	US-10-953-349-1157
9	31	50.8	297	7	US-11-293-697-4132
10	31	50.8	315	6	US-10-953-349-9204
11	31	50.8	369	6	US-10-953-349-3915
12	31	50.8	392	6	US-10-196-749-160
13	31	50.8	400	6	US-10-953-349-9203
14	31	50.8	401	6	US-10-953-349-9202
15	31	50.8	557	7	US-11-316-521-39
16	30.5	50.0	574	7	US-11-121-154-49
17	30	49.2	18	7	US-11-122-986-599
18	30	49.2	96	6	US-10-953-349-16838
19	30	49.2	140	6	US-10-953-349-37648
20	30	49.2	143	6	US-10-953-349-37647
21	30	49.2	262	6	US-10-953-349-24927
22	30	49.2	298	6	US-10-953-349-9347
23	30	49.2	310	6	US-10-953-349-24926
24	30	49.2	381	6	US-10-953-349-24925
25	30	49.2	391	6	US-10-953-349-21132

26	30	49.2	403	6	US-10-953-349-21131	Sequence 21131, A
27	30	49.2	428	6	US-10-953-349-21130	Sequence 21130, A
28	30	49.2	447	7	US-11-122-986-167	Sequence 167, App
29	30	49.2	447	7	US-11-122-986-169	Sequence 169, App
30	30	49.2	503	6	US-10-532-868-18	Sequence 18, Appl
31	30	49.2	537	7	US-11-140-450-33	Sequence 33, Appl
32	30	49.2	750	6	US-10-511-937-2413	Sequence 2413, Ap
33	30	49.2	772	7	US-11-121-154-94	Sequence 94, Appl
34	30	49.2	855	7	US-11-247-437-2	Sequence 2, Appl
35	29	47.5	74	6	US-10-953-349-15283	Sequence 15283, A
36	29	47.5	77	6	US-10-953-349-15282	Sequence 15282, A
37	29	47.5	142	6	US-10-953-349-14978	Sequence 14978, A
38	29	47.5	145	6	US-10-953-349-22188	Sequence 22188, A
39	29	47.5	159	6	US-10-953-349-14977	Sequence 14977, A
40	29	47.5	178	7	US-11-293-697-3713	Sequence 3713, Ap
41	29	47.5	182	6	US-10-953-349-15318	Sequence 15318, A
42	29	47.5	189	6	US-10-953-349-14976	Sequence 14976, A
43	29	47.5	241	6	US-10-953-349-23541	Sequence 23541, A
44	29	47.5	248	6	US-10-953-349-23540	Sequence 23540, A
45	29	47.5	263	6	US-10-953-349-23539	Sequence 23539, A

ALIGNMENTS

RESULT 1

US-11-203-828-5
; Sequence 5, Application US/11203828
; Publication No. US20060110390A1
; GENERAL INFORMATION:
; APPLICANT: LEINWARD, LESLIE
; APPLICANT: SUCHAROV, CARMEN
; TITLE OF INVENTION: INHIBITION OF KU AS A TREATMENT FOR CARDIOVASCULAR DISEASES
; FILE REFERENCE: MYOG:58US
; CURRENT APPLICATION NUMBER: US/11/203,828
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/604,435
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-203-828-5

Query Match 54.1%; Score 33; DB 7; Length 608;
Best Local Similarity 45.5%; Pred.No. 35;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 QGVAGNVRVF 12

Db 461 KAIVGNVFY 471

RESULT 2

US-10-953-349-6894
; Sequence 6894, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6894
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

US-10-953-349-6894

Query Match 53.3%; Score 32.5; DB 6; Length 157;
Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 FQG-VAQNVRVF 10
DB 9 FQGTVDNVRV 19

RESULT 3

US-10-953-349-6893
; Sequence 6893, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 6893

; LENGTH: 208

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-953-349-6893

Query Match 53.3%; Score 32.5; DB 6; Length 208;
Best Local Similarity 72.7%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 FQG-VAQNVRVF 10
DB 60 FQGTVDNVRV 70

RESULT 4

US-10-953-349-6892
; Sequence 6892, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 6892

; LENGTH: 263

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-953-349-6892

Query Match 53.3%; Score 32.5; DB 6; Length 263;
Best Local Similarity 72.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 FQG-VAQNVRVF 10
DB 115 FQGTVDNVRV 125

RESULT 5

US-11-242-505A-36
; Sequence 36, Application US/11242505A
; Publication No. US2006009656A1

; GENERAL INFORMATION:

; APPLICANT: Carroll, Joseph M.

; APPLICANT: Healy, Aileen
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; TITLE OF INVENTION: 14395, 14618, 17692, 58874, 252, 304, 1980, 14717, 9941, 19310, c
; FILE REFERENCE: MPI2001-288PIRCPIOMNIM
; CURRENT APPLICATION NUMBER: US/11/242.505A
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US 10/290, 078
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: US 60/347, 949
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/320, 351
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341, 606
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-242-505A-36

Query Match 52.5%; Score 32; DB 7; Length 980;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQVQAQNVRF 10
DB 434 FQGLSGNERF 443

RESULT 6

US-10-953-349-1159
; Sequence 1159, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 1159

; LENGTH: 164

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-953-349-1159

Query Match 50.8%; Score 31; DB 6; Length 164;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQVQAQNVRFVF 12
DB 82 FEAVDVRVLVVF 93

RESULT 7

US-10-953-349-1158
; Sequence 1158, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 1158
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1158

Query Match 50.8%; Score 31; DB 6; Length 249;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQVQAQNVRFV 12
|:| ||||
Db 167 FEAVDVRVLVF 178

RESULT 8

US-10-953-349-1157
; Sequence 1157, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1157
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-1157

Query Match 50.8%; Score 31; DB 6; Length 250;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQVQAQNVRFV 12
|:| ||||
Db 168 FEAVDVRVLVF 179

RESULT 9

US-11-293-697-4132
; Sequence 4132, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4132
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4132

Query Match 50.8%; Score 31; DB 7; Length 297;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVAQNVRFV 12
||| ||||
Db 234 GVVNVRFV 243

RESULT 10

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME

US-10-953-349-9204
; Sequence 9204, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9204
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9204

Query Match 50.8%; Score 31; DB 6; Length 315;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQVQAQNVRFV 12
|:| ||||
Db 82 FEAVDVRVLVF 93

RESULT 11

US-10-953-349-3915
; Sequence 3915, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3915
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3915

Query Match 50.8%; Score 31; DB 6; Length 369;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQAQNVRFV 11
:||||: |||:
Db 246 RGVATSVRFL 255

RESULT 12

US-10-196-749-160
; Sequence 160, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME

```
; FILE REFERENCE: P3430RIC340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 160
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-160
```

```
Query Match 50.8%; Score 31; DB 6; Length 392;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 FQGVQNVRF 10
|:|:|:|
Db 271 FRGVAASFRF 280
```

```
RESULT 13
US-10-953-349-9203
; Sequence 9203, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9203
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9203
```

```
Query Match 50.8%; Score 31; DB 6; Length 400;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 FQGVQNVRF 12
|:|:|:|
Db 167 FEAVVDRVRLVF 178
```

```
RESULT 14
US-10-953-349-9202
; Sequence 9202, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
```

```
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9202
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9202
```

```
Query Match 50.8%; Score 31; DB 6; Length 401;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 FQGVQNVRF 12
|:|:|:|
Db 168 FEAVVDRVRLVF 179
```

```
RESULT 15
US-11-316-521-39
; Sequence 39, Application US/11316521
; Publication No. US20060111848A1
; GENERAL INFORMATION:
; APPLICANT: Carlow, Clotilde K.S.
; APPLICANT: Foster, Jeremy
; APPLICANT: Zhang, Yinhua
; APPLICANT: Kumar, Sanjay
; TITLE OF INVENTION: Identification and Use of Cofactor Independent Phosphoglycerate
; TITLE OF INVENTION: Mutase as a Drug Target for Pathogenic Organisms and Treatment of
; TITLE OF INVENTION: the Same
; FILE REFERENCE: NEB-230-PCIP-US
; CURRENT APPLICATION NUMBER: US/11/316,521
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 60/483,566
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: PCT/US2004/018200
; PRIOR FILING DATE: 2004-06-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-316-521-39
```

```
Query Match 50.8%; Score 31; DB 7; Length 557;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 GVAQNVRF 10
|:|:|:|
Db 513 GLAQGVRF 520
```

Search completed: June 6, 2006, 00:12:56
Job time : 3.82414 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-28

Perfect score: 59

Sequence: 1 FQGVLRVRFVA 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_8.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

9: Geneseqp2005s.*

10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	59	100.0	12	4	AAB35366 Alpha3bet
2	55	93.2	11	4	AAB35360 Alpha3bet
3	55	93.2	12	4	AAB35352 Alpha3bet
4	55	93.2	12	4	AAB35378 Alpha3bet
5	55	93.2	12	6	ABG72834 Thrombos
6	55	93.2	240	8	ADL70641 Human thr
7	55	93.2	432	8	ADQ39359 Human myo
8	55	93.2	432	8	ADQ39357 Human myo
9	55	93.2	459	4	AAU02916 Angiotens
10	55	93.2	466	3	AAB43602 Human can
11	55	93.2	546	4	AAU02915 Angiotens
12	55	93.2	548	7	ADN02474 TSF polyp
13	55	93.2	555	4	AAU02914 Angiotens
14	55	93.2	731	4	AAU02913 Angiotens
15	55	93.2	1152	3	AAB00042 Human thr
16	55	93.2	1152	5	AAU74771 Human thr
17	55	93.2	1152	5	ABB82285 Human thr
18	55	93.2	1170	4	AAB74450 Human var
19	55	93.2	1170	4	AAB90800 Human she
20	55	93.2	1170	5	AAE25030 Human thr
21	55	93.2	1170	5	AAU75315 Human thr
22	55	93.2	1170	6	ABP96780 Human COP
23	55	93.2	1170	6	ABU03474 Angiotens

24	55	93.2	1170	6	ABG74673 Human THB
25	55	93.2	1170	6	AAB36228 Human THB
26	55	93.2	1170	7	ABR62059 Human thr
27	55	93.2	1170	7	ADN39852 Cancer/an
28	55	93.2	1170	8	ADJ76124 Marker ge
29	55	93.2	1170	8	ADJ75296 Marker ge
30	55	93.2	1170	8	ADL70639 Human thr
31	55	93.2	1170	8	ADL35874 Human thr
32	55	93.2	1170	8	ADQ26070 Thrombos
33	55	93.2	1170	8	ADP54179 Human PRO
34	55	93.2	1170	8	ADQ39358 Human myo
35	55	93.2	1170	8	ADQ39356 Human myo
36	55	93.2	1170	8	ADQ39355 Human myo
37	55	93.2	1170	9	ADZ21688 Thrombos
38	55	93.2	1170	9	AEH87781 Human thr
39	55	93.2	1170	9	AEH46751 Human thr
40	52	88.1	12	4	AAB35373 Alpha3bet
41	52	88.1	12	4	AAB35381 Alpha3bet
42	51	86.4	10	4	AAB35355 Alpha3bet
43	51	86.4	12	4	AAB35364 Alpha3bet
44	51	86.4	12	4	AAB35374 Alpha3bet
45	51	86.4	12	4	AAB35347 Alpha3bet

ALIGNMENTS

RESULT 1

AAB35366

ID AAB35366 standard; peptide; 12 AA.

XX AC AAB35366;

XX DT 08-MAY-2001 (first entry)

XX DE Alpha3bet integrin binding peptide #31.

XX KW Alpha3bet integrin; angiogenesis; cell proliferation; cancer;

KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;

KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX OS Synthetic.

XX FN WO200105812-A2.

XX PD 25-JAN-2001.

XX PF 12-JUL-2000; 2000WO-US018986.

XX PR 15-JUL-1999; 99US-0144549P.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roberts DD, Krutzsch HC;

XX WPI; 2001-182656/18.

XX New peptides that bind to or are recognized by alpha3-beta1 integrins,

XX useful for inhibiting cell adhesion to extracellular matrix, cell

XX motility and proliferation and for treating rheumatoid arthritis and

XX cancer.

XX Claim 4; Page 34; 84pp; English.

XX The present invention provides a number of peptides which bind to

XX alpha3beta1 integrins. They are useful in the modulation of cell adhesion

XX and motility, and in the treatment of cancer, diabetic retinopathy,

XX rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis

XX and restenosis. The present sequence is an example of one of the peptides

XX of the invention

XX SQ Sequence 12 AA;

```
Query Match      100.0%; Score 59; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
DB 1 FQGVQLQNVRFVA 12

RESULT 2
AAB35360
ID AAB35360 standard; peptide; 11 AA.
XX
AC AAB35360;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #25.
XX
KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
CC New peptides that bind to or are recognized by alpha3-beta1 integrins,
CC useful for inhibiting cell adhesion to extracellular matrix, cell
CC motility and proliferation and for treating rheumatoid arthritis and
CC cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;
Query Match      93.2%; Score 55; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00093;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
DB 1 FQGVQLQNVRFV 11

RESULT 4
AAB35378
ID AAB35378 standard; peptide; 12 AA.
XX
AC AAB35378;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3beta1 integrin binding peptide #43.
XX
KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
```

XX WPI; 2001-182656/18.
DR
XX New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Example 2; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
XX Sequence 12 AA;
SQ

Query Match 93.2%; Score 55; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. NO. 0.00093;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 FQGVQLQNVRFV 11
DB 1 FQGVQLQNVRFV 11

RESULT 5
ABG72834
ID ABG72834 standard; peptide; 12 AA.
XX
AC ABG72834;
XX
DT 24-FEB-2003 (first entry)
XX
DE Thrombospondin-1 sequence containing synthetic peptide.
XX
KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
OS Synthetic.
XX
PN WO200281630-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010535.
XX
PR 06-APR-2001; 2001US-0281994P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Krutzsch HC;
XX
DR WPI; 2003-103329/09.
XX
PT A new diagnosis for cancer other than prostate cancer in a mammal useful
PT to detect cancer including lung cancer, particularly small cell lung
PT cancer and melanoma comprises detecting semenogelin in a sample.
XX
PS Example 1; Page 14; 32pp; English.
XX
CC The invention relates to diagnosing cancer other than prostate cancer in
CC a male mammal, comprising assaying a test sample for increased level of
CC semenogelin, or cancer in a female by assaying for the presence of
CC semenogelin. Administering a semenogelin protein or polypeptide fragment
CC or a semenogelin-specific antibody or active fragment, or a recombinant
CC vector expressing the protein or antibody, is useful for inducing an
CC immune response to a cancer in a mammal, where the cancer is not prostate
CC cancer and semenogelin is a marker. The invention is used to diagnose
CC cancer, particularly of epithelial origin such as lung cancer, papillary

CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
CC (SCLC), or a melanoma. The present sequence represents the amino acid
CC sequence of the thrombospondin-1 sequence containing synthetic peptide
CC which binds to alpha-3-beta-1 integrin
XX
SQ Sequence 12 AA;
XX

Query Match 93.2%; Score 55; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. NO. 0.00093;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 FQGVQLQNVRFV 11
DB 1 FQGVQLQNVRFV 11

RESULT 6
ADL70641
ID ADL70641 standard; protein; 240 AA.
XX
AC ADL70641;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human thrombospondin-1 N-terminal domain.
XX
KW Human; thrombospondin-1; epitope; cancer; diagnosis.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Region 23..32 /note= "Heparin binding region"
FT Region 77..82 /note= "Heparin binding region"
FT Region 151..164 /note= "Fibrinogen binding region"
XX
PN WO2004018995-A2.
XX
PD 04-MAR-2004.
XX
PF 20-AUG-2003; 2003WO-US026023.
XX
PR 23-AUG-2002; 2002US-0405494P.
PR 21-APR-2003; 2003US-00419462.
XX
PA (WILL/) WILLIAMS K J.
XX
PI Williams KJ;
XX
DR WPI; 2004-226901/21.
XX
PT New purified thrombospondin fragment extracted from a body fluid, useful
PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
PT or leukemia or as calibrators, indicators, immunogens and analytes.
XX
PS Disclosure; SEQ ID NO 40; 76pp; English.
XX
CC The present sequence is that of the N-terminal domain of human
CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
CC clinical assays for cancer and for generation of antibodies and other
CC binding agents. A method that distinguishes TSP from a TSP fragment or
CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
CC or portion as a target for a binding molecule, e.g. an antibody, to
CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
CC epitope present in TSP but not in the fragment or portion to obtain a
CC quantitation of TSP only; and (3) using the difference between (1) and
CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
CC the TSP fragment or portion is performed in order to detect the presence,
CC or monitor the course, of a disease or condition selected from cancer,

renal failure, renal disease, atopic dermatitis, vasculitis, acute
vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
thrombocytopenic purpura, purpura, an inflammatory condition, a condition
associated with clotting, a condition associated with platelet
activation, a condition associated with intravascular platelet
activation, a condition associated with consumption of platelets, heparin
-induced thrombocytopenia, disseminated intravascular coagulation,
intravascular coagulation, extravascular coagulation, a condition
associated with endothelial activation, a condition associated with
production and/or release of thrombospondin and/or a thrombospondin
fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
reaction, an aspartame reaction, atopic dermatitis, eczema,
hypersensitivity, scleroderma, conditions associated with plugging of
vessels, a condition associated with a cryofibrinogen, a condition
associated with a cryoglobulin, and a condition associated with an anti-
cardiolipin antibody. The cancer is selected from adenoma,
adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
a cancer with vascular invasion, internal cancer, skin cancer, cancer of
the respiratory system, circulatory system, musculoskeletal system,
muscle, bone, a joint, tendon or ligament, digestive system, liver or
biliary system, pancreas, head, neck, endocrine system, reproductive
system (male or female), genitourinary system, kidney, urinary tract,
sensory system, nervous system, lymphoid organ, blood, a gland, mammary
gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
well-differentiated cancer or a moderately differentiated cancer.

Sequence 240 AA;

Query Match 93.2%; Score 55; DB 8; Length 240;

Best Local Similarity 100.0%; Pred. No. 0.027;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 190 FQGVQLQNVRFV 200

RESULT 7

ADQ39359

ID ADQ39359 standard; protein; 432 AA.

XX ADQ39359;

XX 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX cardiant; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakoubova O;

DR WPI; 2004-533949/51.

XX N-PSDB; ADQ38531.

XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acid.

PS Claim 10; SEQ ID NO 1022; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNPs of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 432 AA;

Query Match 93.2%; Score 55; DB 8; Length 432;

Best Local Similarity 100.0%; Pred. No. 0.051;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 208 FQGVQLQNVRFV 218

RESULT 8

ADQ39357

ID ADQ39357 standard; protein; 432 AA.

XX ADQ39357;

XX 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

XX cardiant; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.
 XX Cargill M, Devlin JJ, Iakubova O;
 XX WPI; 2004-533949/51.
 XX N-PSDB; ADO38529.
 XX Identifying an individual who has an altered risk for developing
 PT myocardial infarction by detecting a single nucleotide polymorphism in
 PT the individual's nucleic acids.
 XX Claim 10; SEQ ID NO 1020; 145pp; English.
 XX The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC sequence represents the protein of a human myocardial infarction-
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.
 XX
 SQ Sequence 432 AA;
 Query Match 93.2%; Score 55; DB 8; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FQGVQLQNVRFV 11
 DB 208 FQGVQLQNVRFV 218
 RESULT 9
 AAU02916
 ID AAU02916 standard; protein; 459 AA.
 XX
 AC AAU02916;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #16.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.

XX Homo sapiens.
 XX WO200136632-A2.
 XX 25-MAY-2001.
 XX 17-NOV-2000; 2000WO-IL000766.
 XX 17-NOV-1999; 99IL-00132978.
 XX 10-DEC-1999; 99IL-00133455.
 XX (COMP-) COMPUGEN LTD.
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX WPI; 2001-336004/35.
 XX N-PSDB; AAS06016.
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX Claim 4; Fig 16; 519pp; English.
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding to the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases,
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 459 AA;
 Query Match 93.2%; Score 55; DB 4; Length 459;
 Best Local Similarity 100.0%; Pred. No. 0.055;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FQGVQLQNVRFV 11
 DB 208 FQGVQLQNVRFV 218
 RESULT 10
 AAB43602
 ID AAB43602 standard; protein; 466 AA.
 XX
 AC AAB43602;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1047.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cystostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
 KW vasotropic; antipeptidic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

OS Homo sapiens.
 PN WO200055350-A1.
 XX 21-SEP-2000.
 PD 08-MAR-2000; 2000WO-US005882.
 PF 12-MAR-1999; 99US-0124270P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 PI WPI; 2000-587533/55.
 XX N-PSDB; AAC77811.
 DR Novel isolated nucleic acids comprising sequences encoding peptides
 DR useful for treating or diagnosing e.g. cancer.
 PT Claim 11; Page 1636-1638; 2352pp; English.
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerary; immunomodulator;
 CC antidabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX SQ Sequence 466 AA;
 Query Match 93.2%; Score 55; DB 3; Length 466;
 Best Local Similarity 100.0%; Pred. No. 0.056;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FQGVQLNVRV 11
 |||||
 Db 261 FQGVQLNVRV 271
 RESULT 11
 AAU02915
 ID AAU02915 standard; protein; 546 AA.
 XX AAU02915;
 AC AAU02915;
 XX 12-SEP-2001 (first entry)
 DT Angiotensin converting enzyme (ACEV) splice variant protein #15.
 DE Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX Homo sapiens.
 OS WO200136632-A2.
 PN 25-MAY-2001.
 PD 17-NOV-2000; 2000WO-IL000766.
 PF 17-NOV-1999; 99IL-00132978.
 PR 10-DEC-1999; 99IL-00133455.
 XX (COMP-) COMPUTEN LTD.
 PA Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI WPI; 2001-336004/35.
 XX N-PSDB; AAS06015.
 DR Novel alternative splicing variants e.g. variant of angiotensin
 DR converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 PS Claim 4; Fig 15; 519pp; English.
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX SQ Sequence 546 AA;
 Query Match 93.2%; Score 55; DB 4; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FQGVQLNVRV 11
 |||||
 Db 208 FQGVQLNVRV 218
 RESULT 12
 ADN02474
 ID ADN02474 standard; protein; 548 AA.
 XX ADN02474;
 AC ADN02474;
 XX 17-JUN-2004 (first entry)
 DT TSF polypeptide.
 DE adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
 XX Homo sapiens.
 KW CN1401387-A.
 XX 12-MAR-2003.

XX PF 21-AUG-2002; 2002CN-00129408.
 XX PR 21-AUG-2002; 2002CN-00129408.
 XX PA (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
 XX PI Han Z, Liu P;
 XX WPI; 2003-469302/45.
 DR N-PSDB; ADN02475.
 XX Tumor suppressing polypeptide TSF and gene therapy vector composition.
 PT Claim 2; SEQ ID NO 1; 13pp; Chinese.
 XX The present invention relates to a novel recombinant adenovirus vector
 CC mediated anti-neoplastic composition is prepared through cloning the cDNA
 CC sequence from the human peripheral blood cell by specific primer and
 CC reverse transcription-polymerase chain reaction (RT-PCR) method for
 CC coding TSF polypeptide, constructing in human embryonic kidney cell 293
 CC by AdEasy system, and packaging and expressing the recombinant adenovirus
 CC vector of TSF. It can suppress the growth and transfer of cancer. The
 CC present sequence represents the TSF polypeptide.
 XX
 XX Sequence 548 AA;
 SQ Query Match 93.2%; Score 55; DB 7; Length 548;
 Best Local Similarity 100.0%; Pred. No. 0.067; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0;
 QY 1 FQGVQLQNVRFV 11
 |||||
 DB 208 FQGVQLQNVRFV 218
 RESULT 13
 AAU02914
 ID AAU02914 standard; protein; 555 AA.
 XX
 AC AAU02914;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW noncardiotoxic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200136632-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL000766.
 XX
 PR 17-NOV-1999; 99IL-00132978.
 PR 10-DEC-1999; 99IL-00133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX WPI; 2001-336004/35.
 DR

DR N-PSDB; AAS06014.
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 XX Claim 4; Fig 14; 519pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, noncardiotoxic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 XX Sequence 555 AA;
 SQ Query Match 93.2%; Score 55; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 0.068; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0;
 QY 1 FQGVQLQNVRFV 11
 |||||
 DB 208 FQGVQLQNVRFV 218
 RESULT 14
 AAU02913
 ID AAU02913 standard; protein; 731 AA.
 XX
 AC AAU02913;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #13.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW noncardiotoxic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200136632-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL000766.
 XX
 PR 17-NOV-1999; 99IL-00132978.
 PR 10-DEC-1999; 99IL-00133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX WPI; 2001-336004/35.
 DR

```

DR N-PSDB; AAS06013.
XX
PT Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 13; 519pp; English.
XX
CC The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroiditic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 731 AA;
Query Match 93.2%; Score 55; DB 4; Length 731;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQLQNVRFV 11
DB 208 FQGVQLQNVRFV 218
|||||
|||||

RESULT 15
AAB00042
ID AAB00042 standard; protein; 1152 AA.
XX
AC AAB00042;
XX
DT 08-NOV-2000 (first entry)
XX
DE Human thrombospondon-1 (TSP-1).
XX
KW TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
KW diabetic retinopathy; corneal graft rejection; glaucoma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 361..416
FT /label= Type 1 repeat region
FT Region 417..473
FT /label= Type 1 repeat region
FT Region 474..530
FT /label= Type 1 repeat region
XX
FN WO200044908-A2.
XX
PD 03-AUG-2000.
XX
PF 01-FEB-2000; 2000WO-US002482.
XX
PR 01-FEB-1999; 99US-0118053P.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Lawler JW;
XX
DR WPI; 2000-514823/46.

```

```

XX
PT Nucleic acids encoding chimeric proteins such as cartilage oligomeric
PT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
PT inhibiting angiogenesis and treating diseases such as cancer.
XX
PS Disclosure; Fig 1; 40pp; English.
XX
CC New nucleic acids are described which encode a protein comprising the
CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
CC the TGF (transforming growth factor)-beta activation region of human TSP-
CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
CC tumours in mice models. Thus the nucleic acids and proteins may be useful
CC for treating angiogenesis related diseases such as cancer (by reducing
CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
CC for treating human immunodeficiency virus (HIV) infection. Anti-
CC angiogenic therapy has little toxicity, does not require the therapeutic
CC agent to enter tumour cells or cross the blood-brain barrier, controls
CC tumour growth independently of growth of tumour cell heterogeneity, and
CC does not induce drug resistance
XX
SQ Sequence 1152 AA;
Query Match 93.2%; Score 55; DB 3; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQLQNVRFV 11
DB 190 FQGVQLQNVRFV 200
|||||
|||||

Search completed: June 5, 2006, 22:24:59
Job time : 91.1379 secs

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-28

Perfect score: 59

Sequence: 1 FQGVQLQNVREFA 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	93.2	229	2	thrombospondin 1 -
2	55	93.2	1170	1	thrombospondin 1 p
3	55	93.2	1170	2	thrombospondin 1 p
4	38	64.4	295	2	probable lipolic ac
5	38	64.4	467	2	H+-transporting tw
6	37	62.7	946	2	Ca2+-transporting
7	37	62.7	1020	2	protein envelope C
8	37	62.7	1020	2	Ca2+-transporting
9	37	62.7	1020	2	Ca2+-transporting
10	37	62.7	1172	1	thrombospondin 2 p
11	37	62.7	1172	2	thrombospondin 2 p
12	36	61.0	186	2	probable pilin, ty
13	36	61.0	247	1	probable 3-oxoacyl
14	36	61.0	247	2	3-oxoacyl-[acyl-ca
15	36	61.0	454	2	drug-export protei
16	36	61.0	1054	2	reverse gyrase (co
17	35.5	60.2	218	2	polysialic acid tr
18	35.5	60.2	224	2	kpsr protein - Bsc
19	35	59.3	102	2	hypothetical prote
20	35	59.3	163	2	ribosomal protein
21	35	59.3	248	2	hypothetical prote
22	35	59.3	304	2	hypothetical prote
23	35	59.3	498	2	hypothetical prote
24	35	59.3	740	2	neuraminidase, pro
25	35	59.3	747	2	two component resp
26	35	59.3	759	2	probable autotrans
27	35	59.3	783	2	probable transcrip
28	35	59.3	1178	1	thrombospondin pre
29	34	57.6	82	2	hypothetical prote

30	34	57.6	212	2	D81929	probable imidazole
31	34	57.6	219	2	G95913	Probable cell surf
32	34	57.6	459	2	B95171	NADH oxidase SFl46
33	34	57.6	459	2	B98037	NADH oxidase (EC 1
34	34	57.6	465	2	A84548	probable protein k
35	34	57.6	715	2	S70397	zona pellucida gly
36	34	57.6	1038	2	T02634	rep protein homolo
37	34	57.6	1308	2	T05178	hypothetical prote
38	34	57.6	1829	2	AE1864	hypothetical prote
39	34	57.6	3587	2	I40486	surfactin syntheta
40	34	57.6	3712	2	S18253	laminin alpha-1 ch
41	33	55.9	93	2	T31048	hypothetical prote
42	33	55.9	268	2	B87397	hypothetical prote
43	33	55.9	298	2	C87403	Fdhd protein limpo
44	33	55.9	417	2	H83708	hypothetical prote
45	33	55.9	457	2	S52206	mesy protein - Leu

ALIGNMENTS

RESULT 1

S57957
thrombospondin 1 - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S57957
R:Lafauillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.
submitted to the EMBL Data Library, July 1995
A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expressi
A:Reference number: S57955
A:Accession: S57957
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <LAF>
A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X95911; NID:g899228; PIDN:CI
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

Query Match 93.2%; Score 55; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVREVF 11

Db 190 FQGVQLQNVRFV 200

RESULT 2

TSHUP1
thrombospondin 1 precursor - human
C:Species: Homo sapiens (man)
C>Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: A26155; A34274; A30140; A25812; A05172; A42927
R:Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple ce
A:Reference number: A26155; MUID:87057617; PMID:2430973
A:Accession: A26155
A:Molecule type: mRNA
A:Residues: 1-1170 <LAW>
A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:g37137; PIDN:CI
A>Note: parts of this sequence, including the amino end of the mature protein, were deter
R:laherty, C.D.; German, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA se
A:Reference number: A34274; MUID:89291870; PMID:2544587
A:Accession: A34274
A:Molecule type: DNA
A:Residues: 1-166 <LAH>
A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835
R:Hennessey, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, I
J. Cell Biol. 108, 729-736, 1989
A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>

A;Cross-references: UNIPARC:UPI0000038AB1; EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:R; Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-397 <KOB>

A;Cross-references: UNIPARC:UPI000016B0CA; GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:

R; Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A051172; MUID:86287276; PMID:3461443

A;Accession: A051172

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-374,'RC' <DIX>

A;Cross-references: UNIPARC:UPI000016B140; GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:R; Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992

A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.

A;Reference number: A42927; MUID:92348511; PMID:1379247

A;Accession: A42927

A;Molecule type: protein

A;Residues: 987-1003 <SUN>

A;Cross-references: UNIPARC:UPI00001742C0

A;Note: Cys-992 is shown to have a free sulfhydryl

C;Genetics:

A;Gene: GDB:TBHS1; TSP1; TSP

A;Cross-references: GDB:120438; OMTM:188060

A;Map position: 15q15-15q15

A;Introns: 23/1

A;Note: the list of introns may be incomplete

C;Complex: homotrimer, disulfide linked

C;Function:

A;Description: participates in cell migration and adhesion, and in platelet aggregation

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1170/Product: thrombospondin 1 #status predicted <MAT>

F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>

F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>

F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>

F;491-547/Domain: thrombospondin type 1 repeat homology <EGF1>

F;551-586/Domain: EGF homology <EGF2>

F;650-689/Domain: EGF homology <EGF3>

F;926-928/Region: cell attachment (R-G-D) motif

F;171-232/Disulfide bonds: #status predicted

F;248,360,708,1067/Binding site: carbohydate (Asn) (covalent) #status predicted

F;270,274/Disulfide bonds: interchain #status predicted

F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

F;1051/Binding site: carbohydate (Asn) (covalent) #status absent

Query Match 93.2%; Score 55; DB 1; Length 1170;

Best Local Similarity 100.0%; Pred. No. 0.014;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLQNVRFV 11

|||||

Db 208 FQGVLQNVRFV 218

RESULT 3

A40558

thrombospondin 1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C;Accession: A40558; A37905; B42587; S68787

R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991

A;Title: Characterization of the murine thrombospondin gene.

A;Reference number: A40558; MUID:92128941; PMID:1774063

A;Accession: A40558

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1170 <LAW>

A;Cross-references: UNIPROT:P35441; UNIPARC:UPI0000028012; GB:M62449; GB:M62450; GB:M62451; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469

J. Biol. Chem. 265, 16691-16698, 1990

A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of

A;Reference number: A37905; MUID:90375546; PMID:2398070

A;Accession: A37905

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-490 <BOR>

A;Cross-references: UNIPARC:UPI000016D076; GB:J05605; GB:J05606; NID:g201991; PIDN:AAA404

R; Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.

J. Biol. Chem. 267, 3274-3281, 1992

A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell

A;Reference number: A42587; MUID:92147683; PMID:1371115

A;Accession: B42587

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1152,'P',1154-1170 <LAH>

A;Cross-references: UNIPARC:UPI0000177A96; GB:M87276

A;Note: sequence extracted from NCBI backbone (NCBIP:81501)

R; Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.

FEBS Lett. 387, 36-41, 1996

A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 a

A;Reference number: S68787; MUID:96234006; PMID:8654563

A;Accession: S68787

A;Molecule type: protein

A;Residues: 19-26,'X',28-37 <CH>

A;Cross-references: UNIPARC:UPI0000177A97

C;Complex: homotrimer, disulfide linked

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vor

C;Keywords: calcium binding; glycoprotein; homotrimer

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1170/Product: thrombospondin 1 #status predicted <MAT>

F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>

F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>

F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>

F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>

F;551-586/Domain: EGF homology <EGF>

F;248,360,708,1067/Binding site: carbohydate (Asn) (covalent) #status predicted

Query Match 93.2%; Score 55; DB 2; Length 1170;

Best Local Similarity 100.0%; Pred. No. 0.014;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLQNVRFV 11

|||||

Db 208 FQGVLQNVRFV 218

RESULT 4

E72462

probable lipolic acid synthetase APE2344 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 31-Dec-2004

C;Accession: E72462

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyperthermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: E72462

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-295 <KAW>

A;Cross-references: UNIPROT:Q9Y9E3; UNIPARC:UPI000005E2E9; DDBJ:AP0000064; NID:g5105945; I

A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2344
 C:Superfamily: lipoyl synthase

Query Match 64.4%; Score 38; DB 2; Length 295;
 Best Local Similarity 66.7%; Pred. No. 8.3;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQVRFA 12
 DB 138 FQGVHRLVA 149
 |||| :|||

RESULT 5

D84938
 H+-transporting two-sector ATPase (EC 3.6.3.14), flagellum-specific [imported] - Buchnera sp.
 C:Species: Buchnera sp.
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
 C:Accession: D84938
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A:Reference number: A84930; MUID:20445173; PMID:10993077
 A:Accession: D84938
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-467 <STO>
 A:Cross-references: UNIPARC:UPI00005E44F; GB:AP000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: fliI; BU076
 C:Superfamily: H(+)-transporting ATP synthase; H+-transporting ATP synthase alpha chain
 C:Keywords: hydrolase

Query Match 64.4%; Score 38; DB 2; Length 467;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQVRF 10
 DB 10 FRGILMNLRF 19
 |||| :|||

RESULT 6

S71168
 Ca2+-transporting ATPase (EC 3.6.3.8) ACA1 precursor - Arabidopsis thaliana
 N:Alternate names: envelope Ca2+-ATPase
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 31-Dec-2004
 C:Accession: S71168; S71167
 R:Huang, L.; Berkelman, T.; Franklin, A.E.; Hoffman, N.E.
 submitted to the EMBL Data Library, December 1992
 A:Description: Molecular cloning and characterization of a chloroplast envelope Ca2+-ATPase
 A:Reference number: S71167
 A:Accession: S71168
 A:Molecule type: DNA
 A:Residues: 1-946 <HUA>
 A:Cross-references: UNIPARC:UPI000016DAFA; EMBL:DI13984; NID:g471088; PIDN:BAA03091.1; PI
 A:Accession: S71167
 A:Molecule type: mRNA
 A:Residues: 1-87, 'S', 89-726, 'I', 728-946 <HUZ>
 A:Cross-references: UNIPARC:UPI000016DBCE; EMBL:DI13983; NID:g493621; PIDN:BAA03090.1; PI
 C:Genetics:
 A:Gene: ACA1
 A:Genome: nuclear
 A:Introns: 59/3; 78/1; 654/3; 707/3; 765/1; 864/3
 C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C:Keywords: chloroplast; hydrolase
 F:1-71/Domain: transit peptide (chloroplast) #status predicted <TNP>
 F:72-946/Product: Ca2+-transporting ATPase #status predicted <WAT>
 F:565-750/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 62.7%; Score 37; DB 2; Length 946;
 Best Local Similarity 58.3%; Pred. No. 47;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQVRFA 12
 DB 880 FKGLKNYVFA 891
 |||| :|||

RESULT 7

D86402
 protein envelope Ca2+-ATPase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
 C:Accession: D86402
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, I.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:1130712
 A:Accession: D86402
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1020 <STO>
 A:Cross-references: UNIPROT:Q37145; UNIPARC:UPI0000162EF1; GB:AE005172; NID:g10998927; P1
 C:Genetics:
 A:Map position: 1
 C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding domain

RESULT 8

T51925
 Ca2+-transporting ATPase (EC 3.6.3.8) PEAL [imported] - Arabidopsis thaliana chloroplast
 N:Alternate names: envelope Ca2+-ATPase
 C:Species: chloroplast Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2004
 C:Accession: T51925
 R:Huang, L.; Berkelman, T.; Franklin, A.E.; Hoffman, N.E.
 Proc. Natl. Acad. Sci. U.S.A. 90, 10066-10070, 1993
 A:Title: Characterization of a gene encoding a Ca(2+)-ATPase-like protein in the plastid
 A:Reference number: Z25866; MUID:94052104; PMID:8234257
 A:Accession: T51925
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1020 <HUA>
 A:Cross-references: UNIPROT:Q37145; UNIPARC:UPI000016D30B; EMBL:L08468; PIDN:AAD10211.1
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Gene: PEAL
 A:Genome: chloroplast
 C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C:Keywords: chloroplast; hydrolase

Query Match 62.7%; Score 37; DB 2; Length 1020;
 Best Local Similarity 58.3%; Pred. No. 51;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQVRFA 12
 |||| :|||

Db 954 FKGLKNVYFVA 965

RESULT 9

Ca2+-transporting ATPase (EC 3.6.3.8) PEAL [imported] - Arabidopsis thaliana chloroplast T51926
 N:Alternate names: envelope Ca2+-ATPase
 C:Species: chloroplast Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2004
 C:Accession: T51926
 R:Huang, L.; Berkelman, T.; Franklin, A.E.; Hoffman, N.B.
 Proc. Natl. Acad. Sci. U.S.A. 90, 10066-10070, 1993
 A:Title: Characterization of a gene encoding a Ca(2+)-ATPase-like protein in the plastid
 A:Reference number: Z25866; MUID:94052104; PMID:8234257
 A:Accession: T51926
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1020 <HUA>
 A:Cross-references: UNIPROT:Q37145; UNIPARC:UPI0000125171; EMBL:L08469; PIDN:AAD10212.1
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Gene: PEAL
 A:Genome: chloroplast
 A:Introns: 59/3; 78/1; 728/3; 781/3; 839/1; 938/3
 C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C:Keywords: chloroplast; hydrolase

Query Match 62.7%; Score 37; DB 2; Length 1020;
 Best Local Similarity 58.3%; Pred. No. 51;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12

Db 954 FKGLKNVYFVA 965

RESULT 10

TSHUP2
 thrombospondin 2 precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
 C:Accession: A47379; A42173
 R:LaBell, T.L.; Byers, P.H.
 Genomics 17, 225-229, 1993
 A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: potential role in cell-cell interactions
 A:Reference number: A47379; MUID:94010892; PMID:8406456
 A:Accession: A47379
 A:Molecule type: mRNA
 A:Residues: 1-1172 <LAB>
 A:Cross-references: UNIPROT:P35442; UNIPARC:UPI0000046680; GB:L123350; PIDN:R:LaBell, T.L.; Milewicz, D.J.; Distèche, C.M.; Byers, P.H.
 Genomics 12, 421-429, 1992
 A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression
 A:Reference number: A42173; MUID:92217961; PMID:1559694
 A:Accession: A42173
 A:Molecule type: mRNA
 A:Residues: 560-1172 <LA2>
 A:Cross-references: UNIPARC:UPI00001742C1; GB:M81339
 A:Experimental source: fibroblast
 A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
 C:Genetics:
 A:Gene: GDB:THBS2; TSP2
 A:Cross-references: GDB:128789; OMIM:188061
 A:Map position: 6q27-6q27
 C:Complex: homotrimer, disulfide linked
 C:Function:
 A:Description: participates in cell migration and adhesion, and in platelet aggregation
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type 1 repeat homology
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
 F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>

F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F:553-588/Domain: EGF homology <EGF1>
 F:652-691/Domain: EGF homology <EGF>
 F:928-930/Region: cell attachment (R-G-D) motif
 F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:167-226/Disulfide bonds: #status predicted
 F:266,270/Disulfide bonds: interchain #status predicted
 F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 62.7%; Score 37; DB 1; Length 1172;
 Best Local Similarity 63.6%; Pred. No. 59;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 202 FRGLQLQNVHLV 212

RESULT 11

thrombospondin 2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A42587; A39851
 R:Lahterty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
 J. Biol. Chem. 267, 3274-3281, 1992
 A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
 A:Reference number: A42587; MUID:92147683; PMID:1371115
 A:Accession: A42587
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-1172 <LAH>
 A:Cross-references: UNIPROT:Q03350; UNIPARC:UPI0000029847; GB:L07803; GB:M87275; NID:G34C
 A:Note: sequence extracted from NCBI backbone (NCBIP:81502)
 R:Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
 J. Biol. Chem. 266, 12821-12824, 1991
 A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
 A:Reference number: A39851; MUID:91302287; PMID:1712771
 A:Accession: A39851
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-873 <BOR>
 A:Cross-references: UNIPARC:UPI00016D077; GB:M64866; NID:g201994; PIDN:AAA0432.1; PID:G
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology <VWC>
 F:319-377/Domain: von Willebrand factor type C repeat homology <THR1>
 F:380-431/Domain: thrombospondin type 1 repeat homology <THR2>
 F:436-492/Domain: thrombospondin type 1 repeat homology <THR3>
 F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F:553-588/Domain: EGF homology <EGF1>
 F:652-691/Domain: EGF homology <EGF>

Query Match 62.7%; Score 37; DB 2; Length 1172;
 Best Local Similarity 63.6%; Pred. No. 59;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 202 FRGLQLQNVHLV 212

RESULT 12

B75421
 Probable pilin, type IV - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: B75421
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: B75421

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-186 <WHI>

A;Cross-references: UNIPROT:Q9RUZ7; UNIPARC:UPI00000D3E03; GB:AE001971; GB:AE000513; NID

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR1233

A;Map position: 1

Query Match 61.0%; Score 36; DB 2; Length 186;

Best Local Similarity 87.5%; Pred. No. 13;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQNVRF 9

Db 60 QGVLENRF 67

RESULT 13

A64590

probable 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Helicobacter pylori

C;Species: Helicobacter pylori

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004

C;Accession: A64590

R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: A64590

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-247 <TOM>

A;Cross-references: UNIPROT:Q25286; UNIPARC:UPI00000D3178; GB:AE000570; GB:AE000511; NID

C;Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

C;Keywords: fatty acid biosynthesis; NAD; oxidoreductase

F;6-186/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 61.0%; Score 36; DB 1; Length 247;

Best Local Similarity 50.0%; Pred. No. 17;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRF 10

Db 171 YEGALNRNRF 180

RESULT 14

B71923

3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Helicobacter pylori (strain

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 05-Oct-2004

C;Accession: B71923

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: B71923

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-247 <ARN>

A;Cross-references: UNIPROT:Q9ZLS0; UNIPARC:UPI00000D3648; GB:AE001484; GB:AE001439; NID

A;Experimental source: strain J99

C;Genetics:

A;Gene: fabG

C;Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

C;Keywords: oxidoreductase

F;6-186/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 61.0%; Score 36; DB 2; Length 247;

Best Local Similarity 50.0%; Pred. No. 17;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVRF 10

Db 171 YEGALNRNRF 180

RESULT 15

D86793

drug-export protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C;Accession: D86793

R;Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: D86793

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-454 <STO>

A;Cross-references: UNIPROT:Q9CFW6; UNIPARC:UPI00000C6A19; GB:AE005176; PID:gl2724331; P

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: yniG

C;Superfamily: multidrug-efflux transporter

Query Match 61.0%; Score 36; DB 2; Length 454;

Best Local Similarity 70.0%; Pred. No. 34;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQNVRFVA 12

Db 251 GILKNHRFVA 260

Search completed: June 5, 2006, 22:45:02

Job time : 14.9655 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model
Run on: June 5, 2006, 22:09:41 ; Search time 108.931 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-28
Perfect score: 59
Sequence: 1 FQGVQLNVRFA 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	93.2	229	Q28194_BOVIN	Q28194 bos taurus
2	55	93.2	496	Q7SY84_XENLA	Q7SY84 xenopus lae
3	55	93.2	1170	TSP1_BOVIN	Q28178 bos taurus
4	55	93.2	1170	TSP1_HUMAN	P07996 homo sapien
5	55	93.2	1170	TSP1_MOUSE	P35441 mus musculus
6	55	93.2	1170	Q3TR40_MOUSE	Q3TR40 mus musculus
7	55	93.2	1170	Q71SA3_RAT	Q71SA3 rattus norv
8	55	93.2	1171	Q80YQ1_MOUSE	Q80YQ1 mus musculus
9	55	93.2	1171	Q8CGB2_MOUSE	Q8CGB2 mus musculus
10	55	93.2	1173	TSP1_XENLA	P35448 xenopus lae
11	55	93.2	1225	Q59E99_HUMAN	Q59E99 homo sapien
12	50	84.7	1090	Q5SPG5_BRARE	Q5SPG5 brachydanio
13	50	84.7	1193	Q4S758_TETNG	Q4S758 tetraodon n
14	49	83.1	249	Q5U903_PIG	Q5U903 sus scrofa
15	49	83.1	1171	Q4RLR5_TETNG	Q4RLR5 tetraodon n
16	45	76.3	1034	Q4RQ74_TETNG	Q4RQ74 tetraodon n
17	42	71.2	704	Q4P665_USTMA	Q4P665 ustilago ma
18	42	71.2	727	Q8PRV3_METBA	Q8PRV3 methanosarc
19	42	71.2	730	Q467V1_METBA	Q467V1 methanosarc
20	42	71.2	733	Q8TLX6_METAC	Q8TLX6 methanosarc
21	42	71.2	1549	Q3CJX9_THIET	Q3CJX9 thermoanaer
22	41	69.5	713	Q3F1U8_9BURK	Q3F1U8 burkholderi
23	41	69.5	713	Q44XL2_9BURK	Q44XL2 burkholderi
24	41	69.5	713	Q4LLM8_9BURK	Q4LLM8 burkholderi
25	41	69.5	1168	Q5VH52_CIOIN	Q5VH52 ciona intes
26	40	67.8	647	Q2XGS4_PSEPU	Q2XGS4 pseudomonas
27	40	67.8	647	Q88QB0_PSEPK	Q88QB0 pseudomonas
28	40	67.8	695	Q2LZ43_DROPS	Q2LZ43 drosophila
29	39	66.1	163	R1L0_HAEDU	Q7VKL4 haemophilus
30	39	66.1	189	Q31IL6_THICR	Q31IL6 thiomicrosp
31	39	66.1	212	Q3FA21_NEIGI	Q3FA21 neisseria 9

32	39	66.1	232	2	Q8IS12_9STIC	Q8IS12 oxytricha 1
33	39	66.1	322	2	Q3VHW7_9CHLB	Q3VHW7 pelodictyon
34	39	66.1	327	2	Q5L310_GOKA	Q5L310 geobacillus
35	39	66.1	494	2	Q5AVQ8_EMENI	Q5AVQ8 aspergillus
36	39	66.1	595	2	Q8GNR6_DROME	Q8GNR6 drosophila
37	39	66.1	646	2	Q48HI6_PSEL14	Q48HI6 pseudomonas
38	39	66.1	646	2	Q4ZUA9_PSEU2	Q4ZUA9 pseudomonas
39	39	66.1	646	2	Q883C0_PSESM	Q883C0 pseudomonas
40	39	66.1	721	2	Q9VTH0_DROME	Q9VTH0 drosophila
41	39	66.1	1034	2	Q4DWH6_TRYCR	Q4DWH6 trypanosoma
42	39	66.1	1034	2	Q7Z2B9_TRYCR	Q7Z2B9 trypanosoma
43	39	66.1	1172	2	Q6FXF4_CANGA	Q6FXF4 candida gla
44	39	66.1	1367	2	Q2TVI8_ASPOR	Q2TVI8 aspergillus
45	39	66.1	3242	2	Q859P9_BPN4	Q859P9 bacterioph

ALIGNMENTS

RESULT 1					
Q28194_BOVIN					
ID Q28194_BOVIN	PRELIMINARY;	PRT;	229	AA.	
AC Q28194;					
DT 01-NOV-1996,	integrated into UniProtKB/TrEMBL.				
DT 01-NOV-1996,	sequence version 1.				
DT 07-FEB-2006,	entry version 24.				
DE Thrombospondin-1	(Fragment).				
OS Bos taurus	(Bovine).				
OC Eukaryota;	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia;	Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;				
OC Pecora;	Bovidae; Bovinae; Bos.				
OX NCBI_TaxID=9913;					
RN [1]					
RP NUCLEOTIDE SEQUENCE.					
RX MEDLINE=96331130;	PubMed=8698834;				
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCP19>3.3.CO;2-0;					
RA Lafeuillade B.,	Pellerin S., Keramidas M., Danik M., Chambaz E.M.,				
RA Feige J.J.;					
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced	secreted protein/thrombospondin-2 expression by adrenocorticotrophic				
RT hormone in adrenocortical cells.,"					
RL J. Cell. Physiol. 167:164-172(1996).					
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms					
CC Distributed under the Creative Commons Attribution-NoDerivs License					
EMBL; X89511;	CAA61682.1; -; mRNA.				
DR PIR; S57957;	S57957.				
DR GO; GO:0005198;	P:structural molecule activity; IEA.				
DR GO; GO:0007155;	P:cell adhesion; IEA.				
DR InterPro; IPR013320;	Cona-like subgrp.				
DR InterPro; IPR003129;	Laminin_G_TSP_N.				
DR SMART; SM00210;	TSPN; 1.				
FT NON_TER	1				
FT NON_TER	229				
SQ SEQUENCE	229	AA;	25015	MW;	90D9EBC4E6B669C CRC64;
Query Match 93.2%; Score 55; DB 2; Length 229;					
Best Local Similarity 100.0%; Pred.No. 0.03;					
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	FQGVQLNVRFA	11		
Db	190	FQGVQLNVRFA	200		
RESULT 2					
Q7SY84_XENLA					
ID Q7SY84_XENLA	PRELIMINARY;	PRT;	496	AA.	
AC Q7SY84;					
DT 01-OCT-2003,	integrated into UniProtKB/TrEMBL.				
DT 01-OCT-2003,	sequence version 1.				
DT 07-FEB-2006,	entry version 12.				

DR Pfam; PF05735; TSP_C; 1.
 DR Pfam; PF00093; WVC_1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; WVC; 1.
 DR DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR DR PROSITE; PS01186; EGF_2; 1.
 DR DR PROSITE; PS00026; EGF_3; 2.
 DR DR PROSITE; PS00092; TSP1; 3.
 DR DR PROSITE; PS01208; WVC_1; 1.
 DR DR PROSITE; PS01084; WVC_2; 1.
 KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
 KW Heparin-binding; Repeat; Signal.
 FT SIGNAL 1 18 By similarity.
 FT CHAIN 19 1170 Thrombospondin-1.
 FT 24 221 /FTid=PRO_0000035841.
 FT DOMAIN 316 373 WFC.
 FT DOMAIN 379 429 TSP type-1 1.
 FT DOMAIN 435 490 TSP type-1 2.
 FT DOMAIN 492 547 TSP type-1 3.
 FT DOMAIN 549 587 EGF-like 1.
 FT DOMAIN 588 645 EGF-like 2; calcium-binding (Potential).
 FT DOMAIN 646 690 EGF-like 3.
 FT DOMAIN 723 758 TSP type-3 1.
 FT DOMAIN 759 781 TSP type-3 2.
 FT DOMAIN 782 817 TSP type-3 3.
 FT DOMAIN 818 840 TSP type-3 4.
 FT DOMAIN 841 878 TSP type-3 5.
 FT DOMAIN 879 914 TSP type-3 6.
 FT DOMAIN 915 950 TSP type-3 7.
 FT DOMAIN 951 1170 TSP C-terminal.
 FT REGION 19 232 Heparin-binding (Potential).
 FT MOTIF 926 928 Cell attachment site (Potential).
 FT CARBOHYD 248 248 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 360 360 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 708 708 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1085 1085 N-linked (GlcNAc...) (Potential).
 FT DISULFID 270 270 Interchain (Probable).
 FT DISULFID 274 274 Interchain (Probable).
 FT DISULFID 391 423 By similarity.
 FT DISULFID 395 428 By similarity.
 FT DISULFID 406 413 By similarity.
 FT DISULFID 447 484 By similarity.
 FT DISULFID 451 489 By similarity.
 FT DISULFID 462 474 By similarity.
 FT DISULFID 504 541 By similarity.
 FT DISULFID 508 546 By similarity.
 FT DISULFID 519 531 By similarity.
 FT DISULFID 551 562 By similarity.
 FT DISULFID 556 572 By similarity.
 FT DISULFID 575 586 By similarity.
 FT DISULFID 592 608 By similarity.
 FT DISULFID 599 617 By similarity.
 FT DISULFID 620 644 By similarity.
 FT DISULFID 650 663 By similarity.
 FT DISULFID 657 676 By similarity.
 FT DISULFID 678 689 By similarity.
 FT DISULFID 705 713 By similarity.
 FT DISULFID 718 738 By similarity.
 FT DISULFID 754 774 By similarity.
 FT DISULFID 777 797 By similarity.
 FT DISULFID 813 833 By similarity.
 FT DISULFID 836 856 By similarity.
 FT DISULFID 874 894 By similarity.
 FT DISULFID 910 930 By similarity.
 FT DISULFID 946 1167 By similarity.
 FT CONFLICT 805 805 S -> G (in Ref. 2).
 FT SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3B5FA031A CRC64;

Query Match 93.2%; Score 55; DB 1; Length 1170;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
 |||||
 DB 208 FQGVQLQNVRFV 218

RESULT 4
 TSP1_HUMAN STANDARD; PRT; 1170 AA.
 ID AC P07996; Q15667;
 DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
 DT 01-AUG-1988, sequence version 1.
 DT 07-MAR-2006, entry version 78.
 DE Thrombospondin-1 precursor.
 GN Name=THBS1; Synonyms=TSP, TSP1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 NC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Endothelial cell;
 RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
 RA Lawler J., Hynes R.O.;
 RT "The structure of human thrombospondin, an adhesive glycoprotein with
 RT multiple calcium-binding sites and homologues with several different
 RT proteins.";
 RL J. Cell Biol. 103:1635-1648(1986).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
 RA Hennessy S.W., Frazier W.A., Kim D.D., Deckwerth T.L.,
 RA Baumgartel D.M., Rotwein P., Frazier W.A.;
 RT "Complete thrombospondin mRNA sequence includes potential regulatory
 RT sites in the 3' untranslated region.";
 RL J. Cell Biol. 108:729-736(1989).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 1-397.
 RX MEDLINE=87157592; PubMed=3030396;
 RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
 RT "Partial amino acid sequence of human thrombospondin as determined by
 RT analysis of cDNA clones: homology to malarial circumsporozoite
 RT proteins.";
 RL Biochemistry 25:8418-8425(1986).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 1-374.
 RX MEDLINE=86287276; PubMed=3461443;
 RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
 RT "Characterization of a cDNA encoding the heparin and collagen binding
 RT domains of human thrombospondin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
 RN [5]
 RP NUCLEOTIDE SEQUENCE OF 1-166.
 RX MEDLINE=89291870; PubMed=2544587;
 RA Laherty C.D., Gierman T.M., Dixit V.M.;
 RT "Characterization of the promoter region of the human thrombospondin
 RT gene. DNA sequences within the first intron increase transcription.";
 RL J. Biol. Chem. 264:11222-11227(1989).
 RN [6]
 RP NUCLEOTIDE SEQUENCE OF 1028-1170.
 RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
 RT "Expression of thrombospondin in chronic inflammation: neutrophils
 RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
 RP THR-450; TRP-498 AND THR-507.
 RP TISSUE=platelet;
 RC MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;

RA Hofsteenge J., Hwuiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
module."; J. Biol. Chem. 276:6485-6498(2001).
RL [8]
RN THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RP MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RX Huwiler K.G., Vesling M.M., Anis D.S., Mosher D.F.;
RA "Biophysical characterization, including disulfide bond assignments,
of the anti-angiogenic type 1 domains of human thrombospondin-1."; J.
RL Biochemistry 41:14329-14339(2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
hydrazide chemistry, and mass spectrometry."; J.
RL Proteome Res. 4:2070-2080(2005).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
laminin, type V collagen and integrins alpha-V/beta-1, alpha-
V/beta-3 and alpha-11b/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04655; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAG61237.1; ALT_SEQ; mRNA.
DR EMBL; J04835; AAG61178.1; -; Genomic_DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1Z44; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; Gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF_.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWFC; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
KW 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Calcium; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT FTIG=PRO 000035842.
FT TSP N-terminal.
FT DOMAIN 24 221 VWFC.
FT DOMAIN 316 373 TSP type-1 1.
FT DOMAIN 379 429 TSP type-1 2.
FT DOMAIN 435 490 TSP type-1 3.
FT DOMAIN 492 547 EGF-like 1.
FT DOMAIN 549 587 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 588 645 EGF-like 3.
FT DOMAIN 646 690 TSP type-3 1.
FT DOMAIN 723 758 TSP type-3 2.
FT DOMAIN 759 781 TSP type-3 3.
FT DOMAIN 782 817 TSP type-3 4.
FT DOMAIN 818 840 TSP type-3 5.
FT DOMAIN 841 878 TSP type-3 6.
FT DOMAIN 879 914 TSP type-3 7.
FT DOMAIN 915 950 TSP C-terminal.
FT DOMAIN 951 1170 Heparin-binding (Potential).
FT MOTIF 926 928 Cell attachment site (Potential).
FT CARBOHYD 248 248 N-linked (GlcNAc...).(Potential).
FT CARBOHYD 360 360 N-linked (GlcNAc...).(Potential).
FT CARBOHYD 385 385 C-linked (Man).
FT CARBOHYD 394 394 /FTIG=CAR 000205.
FT CARBOHYD 438 438 /FTIG=CAR 000206.
FT CARBOHYD 441 441 /FTIG=CAR 000207.
FT CARBOHYD 450 450 /FTIG=CAR 000208.
FT CARBOHYD 450 450 /FTIG=CAR 000209.
FT CARBOHYD 498 498 /FTIG=CAR 000210.
FT CARBOHYD 507 507 /FTIG=CAR 000211.
FT CARBOHYD 708 708 N-linked (GlcNAc...).(Potential).
FT CARBOHYD 1067 1067 N-linked (GlcNAc...).(Potential).
FT DISULFID 270 270 Interchain (Probable).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 391 423 Interchain (Probable).
FT DISULFID 395 428 Interchain (Probable).
FT DISULFID 406 413 Interchain (Probable).
FT DISULFID 447 484 Interchain (Probable).
FT DISULFID 451 489 Interchain (Probable).
FT DISULFID 462 474 Interchain (Probable).
FT DISULFID 504 541 Interchain (Probable).
FT DISULFID 508 546 Interchain (Probable).

Query Match 93.2%; Score 55; DB 1; Length 1170;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNRV 11

Db 208 FQGVQLQNRV 218

RESULT 5

TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 57.
DE Thrombospondin-1 precursor.
GN Name=Thbs1; Synonyms=Tspl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-490.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfi D., Devarayalu S., Franson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
RN [4]
RP PROTEIN SEQUENCE OF 19-37.
RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
RT "Expression and initial characterization of recombinant mouse thrombospondin 1 and thrombospondin 3.";
RL FEBS Lett. 387:36-41(1996).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibronectin, fibronectin, laminin, type V collagen and integrins alpha-V/beta-1, alpha-V/beta-3 and alpha-ITB/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M62470; AAA50611.1; -; Genomic DNA.
DR EMBL; M62450; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62451; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62452; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62453; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62454; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62455; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62456; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62457; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62458; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62459; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62460; AAA50611.1; JOINED; Genomic DNA.

DR EMBL; M62461; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62462; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62463; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62464; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62465; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62466; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62467; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62468; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62469; AAA50611.1; JOINED; Genomic DNA.
DR EMBL; M62476; AAA50611.1; -; mRNA.
DR EMBL; J05606; AAA40431.1; -; Genomic DNA.
DR EMBL; J05605; AAA40431.1; JOINED; Genomic DNA.
DR PIR; A40558; A40558.
DR HSSP; P07996; ILSL.
DR SWR; P35441; 549-1169.
DR Ensembl; ENSMUSG00000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008085; tsp_1.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00266; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
DR Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
KW Glycoprotein; Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO_0000035843.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT CARBOHYD 708 708

```

FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 270 Interchain (Probable).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 391 423 By similarity.
FT DISULFID 395 428 By similarity.
FT DISULFID 406 413 By similarity.
FT DISULFID 447 484 By similarity.
FT DISULFID 451 489 By similarity.
FT DISULFID 462 474 By similarity.
FT DISULFID 504 541 By similarity.
FT DISULFID 508 546 By similarity.
FT DISULFID 519 531 By similarity.
FT DISULFID 551 562 By similarity.
FT DISULFID 556 572 By similarity.
FT DISULFID 575 586 By similarity.
FT DISULFID 592 608 By similarity.
FT DISULFID 599 617 By similarity.
FT DISULFID 620 644 By similarity.
FT DISULFID 650 663 By similarity.
FT DISULFID 657 676 By similarity.
FT DISULFID 678 689 By similarity.
FT DISULFID 705 713 By similarity.
FT DISULFID 718 738 By similarity.
FT DISULFID 754 774 By similarity.
FT DISULFID 777 797 By similarity.
FT DISULFID 813 833 By similarity.
FT DISULFID 836 856 By similarity.
FT DISULFID 874 894 By similarity.
FT DISULFID 910 930 By similarity.
FT DISULFID 946 1167 By similarity.
FT CONFLICT 1025 1025 F -> L (in Ref. 2).
SQ SEQUENCE 1170 AA; 129647 MW; 0443B493615E7F06 CRC64;

Query Match 93.2%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. NO. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQVRV 11
Db 208 FQGVQLQVRV 218

RESULT 6
Q3TR40_MOUSE PRELIMINARY; PRT; 1170 AA.
AC Q3TR40;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 5.
DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DE clone:A330055N06 product:thrombospondin 1, full insert sequence.
GN Name=rhbe1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX RAVASCI T., REED J.C., REID D.J., REID J., RING B.Z., RINGWALD M.,
RA CARNINCI P., HAYASHIZAKI Y.;
RA "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RC PubMed=16141072; DOI=10.1126/science.1112014;
RX CARNINCI P., KASUKAWA T., KATAYAMA S., GOUGH J., FRITH M.C., MAEDA N.,
RA CYAMA R., RAVASI I., LENHARD B., WELLS C., KODZIUS R., SHIMOKAWA K.,
RA BAJIC V.B., BRENNER S.E., BATALOV S., FORREST A.K., ZAVOLAN M.,
RA DAVIS M.J., WILMING L.G., AIDINIS V., ALLEN J.E.,
RA AMBESI-IMPIOMBATO A., APWEILER R., ATURALIYA R.N., BAILEY T.L.,

```

```

RA BANSAL M., BAXTER L., BEISEL K.W., BERSANO T., BONO H., CHALK A.M.,
RA CHIU K.P., CHOUDHARY V., CHRISTOFFELS A., CLUTTERBUCK D.R.,
RA CROWE M.L., DALLA E., DALRYMPLE B.P., DE BONO B., DELLA GATTA G.,
RA DI BERNARDO D., DOWN T., ENGSTROM P., FAGIOLINI M., FAULKNER G.,
RA FLETCHER C.F., FUKUSHIMA T., FURUNO M., FUTAKI S., GARIBOLDI M.,
RA GEORGII-HEMMING P., GINGERAS T.R., GOJOBORI T., GREEN R.E.,
RA GUSTINCICH S., HARBERS M., HAYASHI Y., HENSCH T.K., HIROKAWA N.,
RA HILL D., HUMINIECKI L., IACONO M., IKEO K., IWAMA A., ISHIKAWA T.,
RA JAKT M., KANAPIN A., KATOH M., KAWASAWA Y., KESLO J., KITAMURA H.,
RA KITANO H., KOLLIAS G., KRISHNAN S.P., KRUGER A., KUMMERFELD S.K.,
RA KUROCHKIN I.V., LAREAU L.F., LAZAREVIC D., LIPOVICH L., LIU J.,
RA LIUNI S., MCWILLIAM S., MADAN BABU M., MADERA M., MARCHIONNI L.,
RA MATSUDA H., MATSUZAWA S., MIKI H., MIGNONE F., MIYAKE S., MORRIS K.,
RA MOTTAGUI-TABAR S., MULDER N., NAKANO N., NAKAUCHI H., NG P.,
RA NILSSON R., NISHIGUCHI S., NISHIKAWA S., NORI F., OHARA O.,
RA OKAZAKI Y., ORLANDO V., PANG K.C., PAVAN W.J., PAVESI G., PESOLE G.,
RA PETROVSKY N., PIAZZA S., REED J., REID J.P., RING B.Z., RINGWALD M.,
RA ROST B., RUAN Y., SALZBERG S.L., SANDELIN A., SCHNEIDER C.,
RA SCHONBACH C., SEKIGUCHI K., SEMPLE C.A., SENO S., SESSE L.,
RA SHIBATA Y., SHIMADA H., SHIMADA K., SILVA D., SINCLAIR B.,
RA SPERLING S., STUPKA E., SUGIURA K., SULTANA R., TAKENAKA Y., TAKI K.,
RA TAMMOJA K., TAN S.L., TANG S., TAYLOR M.S., TEGNER J., TEICHMANN S.A.,
RA UEDA H.R., VAN NIMWEGEN E., VERARDO R., WEI C.L., YAGI K.,
RA YAMANISHI H., ZABAROVSKY E., ZHU S., ZIMMER A., HIDE W., BULT C.,
RA GRIMMOND S.M., TEASDALE R.D., LIU E.T., BRUSIC V., QUACKENBUSH J.,
RA WAHLESTEDT C., MATTICK J.S., HUME D.A., KAI C., SASAKI D., TOMARU Y.,
RA FUKUDA S., KANAMORI-KATAYAMA M., SUZUKI M., AOKI J., ARAKAWA T.,
RA IIDA J., INAMURA K., ITOH M., KATO T., KAWAJI H., KAWAGASHIRA N.,
RA KAWASHIMA T., KOJIMA M., KONDO S., KONNO H., NAKANO K., NINOMIYA N.,
RA NISHIO T., OKADA M., PLESSY C., SHIBATA K., SHIRAKI T., SUZUKI S.,
RA TAGAMI M., WAKI K., WATAHIKI A., OKAMURA-OHO Y., SUZUKI H., KAWAI J.,
RA HAYASHIZAKI Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA OKAZAKI Y., FURUNO M., KASUKAWA T., ADACHI J., BONO H., KONDO S.,
RA NIKALDO I., OSAITO N., SAITO R., SUZUKI H., YAMANAKA I., KIYOSAWA H.,
RA YAGI K., TOMARU Y., HASEGAWA Y., NOGAMI A., SCHONBACH C., GOJOBORI T.,
RA BALDARELLI R., HILL D.P., BULT C., HUME D.A., QUACKENBUSH J.,
RA SCHIRML L.M., KANAPIN A., MATSUDA H., BATALOV S., BEISEL K.W.,
RA BLAKE J.A., BRADT D., BRUSIC V., CHOITHIA C., CORBANI L.E., COUSINS S.,
RA DALLA E., DRAGANI T.A., FLETCHER C.F., FORREST A., FRAZER K.S.,
RA GAESTERLAND T., GARIBOLDI M., GISSI C., GODZIK A., GOUGH J.,
RA GRIMMOND S., GUSTINCICH S., HIROKAWA N., JACKSON I.J., JARVIS E.D.,
RA KANAI A., KAWAJI H., KAWASAWA Y., KEDZIERSKI R.M., KING B.L.,
RA KONAGAYA A., KUROCHKIN I.V., LEE Y., LENHARD B., LYONS P.A.,
RA MAGLOTT D.R., MALTAIS L., MARCHIONNI L., MCKENZIE L., MIKI H.,
RA NAGASHIMA T., NUMATA K., OKIDO T., PAVAN W.J., PERTEA G., PESOLE G.,
RA PETROVSKY N., PILLAI R., PONTIUS J.U., QI D., RAMACHANDRAN S.,
RA RAVASI T., REED J.C., REED D.J., REID J., RING B.Z., RINGWALD M.,
RA SANDELIN A., SCHNEIDER C., SEMPLE C.A., SETOU M., SHIMADA K.,
RA SULTANA R., TAKENAKA Y., TAYLOR M.S., TEASDALE R.D., TOMITA M.,
RA VERARDO R., WAGNER L., WAHLESTEDT C., WANG Y., WATANABE Y., WELLS C.,
RA WILMING L.G., WYNHAW-BORIS A., YANAGISAWA M., YANG I., YANG L.,
RA YUAN Z., ZAVOLAN M., ZHU Y., ZIMMER A., CARNINCI P., HAYASU N.,
RA HIROZANE-KISHIKAWA T., KONNO H., NAKAMURA M., SAKAZUME N., SATO K.,
RA SHIRAKI T., WAKI K., KAWAI J., AIZAWA K., ARAKAWA T., FUKUDA S.,
RA HARA A., HASHIZUME W., IMOTANI K., ISHII Y., ITOH M., KAGAWA I.,
RA MIYAZAKI A., SAKAI K., SASAKI D., SHIBATA K., SHINAGAWA A.,
RA YASUNISHI A., YOSHINO M., WATERSTON R., LANDER E.S., ROGERS J.,
RA BIRNEY E., HAYASHIZAKI Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```



```

DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS0026; EGF_3; 2.
DR PROSITE: PS0092; TSP1; 3. UNKNOWN_1.
DR PROSITE: PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE: PS0184; VWFC_2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred.No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNRFV 11
Db 208 FQGVQNRFV 218
|||||

RESULT 8
Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1_
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Warra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

```

CC EMBL: BC050917; AAH50917.1; -, mRNA.
DR HSP: P07996; 1LSL.
DR SMR: Q80YQ1; 835-1170.
DR Ensembl: ENSMUSG00000040152; Mus musculus.
DR MGI: MGI:98737; Thbs1.
DR GO: GO:0005615; C:extracellular space; RCA.
DR GO: GO:0005615; C:extracellular space; IDA.
DR GO: GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro: IPR013320; ConA_like_subgrp.
DR InterPro: IPR006210; EGF.
DR InterPro: IPR000742; EGF_3.
DR InterPro: IPR001881; EGF_Ca bd.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR013032; EGF_like_reg.
DR InterPro: IPR003329; Laminin_G_TSP_N.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR008085; TSP_1.
DR InterPro: IPR003367; tsp_3.
DR InterPro: IPR008859; TSP_C.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS0026; EGF_3; 2.
DR PROSITE: PS0092; TSP1; 3. UNKNOWN_1.
DR PROSITE: PS01208; VWFC_1; UNKNOWN_1.
DR PROSITE: PS0184; VWFC_2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 93.2%; Score 55; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred.No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNRFV 11
Db 208 FQGVQNRFV 218
|||||

RESULT 9
Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q8CGB2_
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 Jyg-MC(B) cDNA, RIKEN full-
length enriched library, clone:G930018021 product:thrombospondin 1,
full insert sequence).
DE full insert sequence.
DE Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Mammary tumor metastasized to lung. Tumor
arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Warra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J.J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor
 RC arose spontaneously;
 RC Strausberg R.;
 RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RX Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 RN [4]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC PubMed=16141072; DOI=10.1126/science.1112014;
 RX Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Adinolfi V., Allen J.E.,
 RA Ambesi-Impiombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Belsel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli N., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J.P., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
 RA Tamaoka K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Tesdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 RN [5]
 RN NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566(2005).
 RN [6]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kiyosawa H.,
 RA Nikaido I., Oatono N., Saito R., Suzuki H., Yamanaka I., Gojobori T.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C.F., Forrest A., Frazer K.S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Gojobori T., Gough J.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kageyama A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Tesdale R.D., Tomita M.,
 RA Verardo R., Wagner W., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Xuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayateu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imortani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [7]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [8]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [9]

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=2053013; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[10]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Mammary gland;
RC Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC [10]
CC EMBL; BC042422; AAH42422.1; -; mRNA.
CC EMBL; AK145202; BAE26293.1; -; mRNA.
CC HSSP; P07996; ILSL.
CC
CC Query Match 93.2%; Score 55; DB 2; Length 1171;
CC Best Local Similarity 100.0%; Pred. No. 0.18; 0; Indels 0; Gaps 0;
CC Matches 11; Conservative 0; Mismatches 0;
CC
CC QY 1 FQGVQLNVRV 11
CC Db 208 FQGVQLNVRV 218
CC
CC RESULT 10
CC TSPI_XENLA STANDARD; PRT; 1173 AA.
CC ID TSPI_XENLA STANDARD; PRT; 1173 AA.
CC AC P3548;
CC DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
CC DT 01-JUN-1994, sequence version 1.
CC DT 07-MAR-2006, entry version 54.
CC DE Thrombospondin-1 precursor.
CC GN Name=thb1; Synonyms=tspl;
CC OS Xenopus laevis (African clawed frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
CC OC Xenopodidae; Xenopus; Xenopus.
CC OX NCBI_TaxID=8355;
CC [1]_TaxID=8355;
CC
CC NUCLEOTIDE SEQUENCE [MRNA].
CC Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
CC RT "Cloning, characterization and expression of thrombospondin-1 in
CC RT Xenopus laevis embryos.";
CC RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.

```

```

CC CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
CC HSSP; P07996; ILSL.
CC SMR; P35448; 552-1172.
CC InterPro; IPR013320; Cona_like_subgrp.
CC InterPro; IPR006210; EGF_3.
CC InterPro; IPR000742; EGF_Ca_bd.
CC InterPro; IPR001881; EGF_like.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR013032; EGF_like_reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSP_C.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS50026; EGF_3; 2.
CC PROSITE; PS50092; TSP1; 3.
CC PROSITE; PS01208; VWFC_1; 1.
CC PROSITE; PS0184; VWFC_2; 1.
CC Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1173 Thrombospondin-1.
FT DOMAIN 23 224 /FTID=PRO_0000035844.
FT DOMAIN 319 376 TSP N-terminal.
FT DOMAIN 382 432 TSP type-1 1.
FT DOMAIN 438 493 TSP type-1 2.
FT DOMAIN 495 550 TSP type-1 3.
FT DOMAIN 550 590 EGF-like 1.
FT DOMAIN 591 648 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 649 693 EGF-like 3.
FT DOMAIN 726 761 TSP type-3 1.
FT DOMAIN 762 784 TSP type-3 2.
FT DOMAIN 785 820 TSP type-3 3.
FT DOMAIN 821 843 TSP type-3 4.
FT DOMAIN 844 881 TSP type-3 5.
FT DOMAIN 882 917 TSP type-3 6.
FT DOMAIN 918 953 TSP type-3 7.
FT DOMAIN 954 1173 TSP C-terminal.
FT REGION 23 235 Heparin-binding (Potential).
FT MOTIF 929 931 Cell attachment site (Potential).
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 363 363 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 705 705 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 711 711 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1070 1070 N-linked (GlcNAc...) (Potential).
FT DISULFID 394 426 By similarity.
FT DISULFID 398 431 By similarity.
FT DISULFID 409 416 By similarity.
FT DISULFID 450 487 By similarity.
FT DISULFID 454 492 By similarity.
FT DISULFID 465 477 By similarity.
FT DISULFID 507 544 By similarity.
FT DISULFID 511 549 By similarity.
FT DISULFID 522 534 By similarity.

```

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
Distributed under the Creative Commons Attribution-NoDerivs License


```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 602 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 93.2%; Score 55; DB 1; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 211 FQGVQLQNVRFV 221

RESULT 11
Q59E99 HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RT Ohara O., Nagase T., Kikuno F.R.;
RL "None Title.";
Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AB209912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939, 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0003198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003387; tsp_3.
DR InterPro; IPR008859; TSP_C.
```

```
DR InterPro; IPR001007; WF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; WF_C_1; 1.
DR PROSITE; PS50184; WF_C_2; 1.
DR Cell adhesion; EGF-like domain.
KW NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16E57157B12 CRC64;

Query Match 93.2%; Score 55; DB 2; Length 1225;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 263 FQGVQLQNVRFV 273

RESULT 12
Q5SPG5 BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n\ (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AL928866; CAI20599.1; -; Genomic_DNA.
DR SMR; Q5SPG5; 751-804, 754-1089.
DR Ensembl; ENSDARG0000010785; Danio rerio.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0003198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
```

```

DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP 1; 2.
DR Pfam; PF02412; TSP 3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS00018; EF_HAND_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF 3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; WFC 1; 1.
DR PROSITE; PS0184; WFC_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 1090;
Best Local Similarity 90.9%; Pred. No. 1.6;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 185 FNGVLQNVRFV 195

RESULT 13
ID Q4S758 TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DE Chromosome 14 SCAP14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00022976001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Oucif-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bieumont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

```

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; CAAE01014723; CAG03524.1; -; Genomic_DNA.
CC SMR; Q4S758; 811-1148.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR00742; EGF 3.
CC InterPro; IPR01881; EGF_Ca bd.
CC InterPro; IPR013032; EGF-like reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR01007; VWF_C.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00090; TSP 1; 3.
CC Pfam; PF02412; TSP 3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC Pfam; PF00093; VMC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VMC; 1.
CC PROSITE; PS01186; EGF 2; UNKNOWN_1.
CC PROSITE; PS00026; EGF 3; 2.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS01208; WFC 1; 1.
CC PROSITE; PS0184; WFC_2; 1.
KW Cell adhesion.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 1193;
Best Local Similarity 90.9%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 187 FNGVLQNVRFV 197

RESULT 14
Q5U903 FIG PRELIMINARY; PRT; 249 AA.
ID Q5U903 FIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang K., Mauc G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AY773342; AAV38110.1; -; mRNA.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR001007; VWF_C.

```

```

DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VMC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
FT NON_TER 249
FT SEQUENCE 249 AA; 27560 MW; 465D664BE0329C32 CRC64;

Query Match 83.1%; Score 49; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFV 11
Db 1 QGVQLQNVRFV 10

RESULT 15
Q4RLR5 TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Chromosome 10 SCAR15019, whole genome shotgun sequence. (Fragment).
OS ORFNames=GSTENG00032374001;
GN Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC
CC EMBL; CAAE01015019; CAG10667.1; -; Genomic_DNA.
CC SMR; Q4RLR5; 834-887, 837-1171.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.

```

```

DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR01881; EGF_Ca_bd.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
KW Cell adhesion.
FT NON_TER 1171
FT SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 83.1%; Score 49; DB 2; Length 1171;
Best Local Similarity 90.9%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 207 FTQVQLQNVRFV 217

Search completed: June 5, 2006, 22:42:43
Job time : 110.931 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-28

Perfect score: 59

Sequence: 1 FQGVQLQNVRFVA 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/pCTUS_COMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	93.2	825	2	US-09-949-002-482
2	55	93.2	831	2	US-09-939-853A-97
3	55	93.2	831	2	US-09-939-853A-98
4	55	93.2	1170	1	US-08-313-288B-20
5	55	93.2	1170	2	US-09-657-472-2
6	55	93.2	1170	2	US-09-949-002-350
7	39	66.1	731	2	US-09-270-767-42057
8	37	62.7	1045	2	US-09-949-016-1112
9	37	62.7	1172	1	US-08-313-288B-19
10	37	62.7	1172	2	US-09-949-016-6333
11	35	59.3	162	2	US-09-732-210-527
12	35	59.3	415	2	US-09-489-039A-10457
13	35	59.3	465	2	US-09-769-787-164
14	34	57.6	169	2	US-09-134-000C-3511
15	34	57.6	459	2	US-09-583-110-3429
16	34	57.6	715	1	US-08-484-993B-10
17	34	57.6	715	1	US-08-484-158B-10
18	34	57.6	715	1	US-08-484-596A-10
19	34	57.6	715	1	US-08-480-150A-10
20	34	57.6	715	2	US-08-458-731-10
21	34	57.6	715	2	US-08-149-223A-10
22	34	57.6	944	2	US-09-134-000C-5578
23	34	57.6	3712	2	US-10-037-417-48
24	34	57.6	3712	2	US-10-037-417-51
25	33	55.9	133	2	US-09-513-999C-7744
26	33	55.9	175	2	US-09-230-637-24

```
27 33 55.9 175 2 US-10-003-632C-5 Sequence 5, Appli
28 33 55.9 390 2 US-09-328-352-4891 Sequence 4891, Ap
29 33 55.9 400 2 US-09-252-991A-20800 Sequence 20800, A
30 33 55.9 446 2 US-09-199-637A-267 Sequence 267, App
31 33 55.9 484 2 US-09-248-796A-16180 Sequence 16180, A
32 33 55.9 538 2 US-09-134-000C-4331 Sequence 4331, Ap
33 33 55.9 743 2 US-09-708-392-1 Sequence 1, Appli
34 33 55.9 750 2 US-09-949-016-6004 Sequence 6004, Ap
35 33 55.9 750 2 US-09-949-016-6791 Sequence 6791, Ap
36 33 55.9 750 2 US-09-949-016-6792 Sequence 6792, Ap
37 33 55.9 750 2 US-09-949-016-6793 Sequence 6793, Ap
38 33 55.9 750 2 US-09-948-429C-63 Sequence 63, Appl
39 33 55.9 751 2 US-09-949-016-9533 Sequence 9533, Ap
40 33 55.9 751 2 US-09-949-016-9534 Sequence 9534, Ap
41 33 55.9 751 2 US-09-949-016-9535 Sequence 9535, Ap
42 33 55.9 751 2 US-09-949-016-9536 Sequence 9536, Ap
43 33 55.9 751 2 US-09-949-016-9537 Sequence 9537, Ap
44 33 55.9 751 2 US-09-949-016-9538 Sequence 9538, Ap
45 33 55.9 751 2 US-09-949-016-9539 Sequence 9539, Ap
```

ALIGNMENTS

RESULT 1

```
US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-482
```

```
Query Match 93.2%; Score 55; DB 2; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.027; 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;
```

```
QY 1 FQGVQLQNVRFV 11
Db 308 FQGVQLQNVRFV 318
```

RESULT 2

```
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
```

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match
Best Local Similarity 93.2%; Score 55; DB 2; Length 831;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
Db 208 FQGVQNVRV 218

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match
Best Local Similarity 93.2%; Score 55; DB 2; Length 831;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
Db 208 FQGVQNVRV 218

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match
Best Local Similarity 93.2%; Score 55; DB 2; Length 831;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
Db 208 FQGVQNVRV 218

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match
Best Local Similarity 93.2%; Score 55; DB 2; Length 1170;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
Db 208 FQGVQNVRV 218

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
; US-09-949-002-350

Query Match          93.2%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQLQVRVF 11
Db      208 FQGVQLQVRVF 218
      |||||
RESULT 7
US-09-270-767-42057
; Sequence 42057, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42057
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-42057

Query Match          66.1%; Score 39; DB 2; Length 731;
Best Local Similarity 80.0%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 QGVQLQVRVF 11
Db      238 QGVQLQVRVF 247
      |||||
RESULT 8
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-11112

Query Match          62.7%; Score 37; DB 2; Length 1045;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 FQGVQLQVRVF 11
Db      281 FQGVQLQVRVF 291
      |||||
RESULT 9
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-288B-19

Query Match          62.7%; Score 37; DB 1; Length 1172;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 FQGVQLQVRVF 11
Db      202 FQGVQLQVRVF 212
      |||||
RESULT 10
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6333

Query Match 62.7%; Score 37; DB 2; Length 1172;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
|:|:|:|:|
DB 202 FRGLQLQNVHLV 212

RESULT 11
US-09-732-210-527
; Sequence 527, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 527
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-732-210-527

Query Match 59.3%; Score 35; DB 2; Length 162;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:|:|:|:|
DB 112 FEGKIQDVFLA 123

RESULT 12
US-09-489-039A-10457
; Sequence 10457, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10457
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10457

Query Match 59.3%; Score 35; DB 2; Length 415;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VLQNVRFVA 12
|:|:|:|:|
DB 220 VLKMLRFVA 228

RESULT 13
US-09-769-787-164
; Sequence 164, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 164
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-164

Query Match 59.3%; Score 35; DB 2; Length 465;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRF 10
|:|:|:|:|
DB 245 FNGTLENIKF 254

RESULT 14
US-09-134-000C-3511
; Sequence 3511, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3511
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (29)..(70)

; OTHER INFORMATION: Amino acids 29, 30, 59 & 70 are Xaa wherein Xaa = any amino acid.
US-09-134-000C-3511

Query Match 57.6%; Score 34; DB 2; Length 169;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FQGVLRVVR 9
DB 76 FQGVLRVVR 84

RESULT 15

US-09-583-110-3429
; Sequence 3429, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3429
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3429

Query Match 57.6%; Score 34; DB 2; Length 459;
Best Local Similarity 54.5%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 FQGVLRVVR 11
DB 132 FQGVLRVVR 142

Search completed: June 5, 2006, 22:48:55
Job time : 24.8966 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-28
Perfect score: 59
Sequence: 1 FQGVQLQNVRFVA 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	93.2	12	4	US-10-474-213-28
2	55	93.2	240	4	US-10-419-462-40
3	55	93.2	240	5	US-10-782-968-40
4	55	93.2	432	5	US-10-741-600-1020
5	55	93.2	432	5	US-10-741-600-1022
6	55	93.2	459	6	US-11-043-806-462
7	55	93.2	466	3	US-09-925-301-1047
8	55	93.2	555	6	US-11-043-806-454
9	55	93.2	578	6	US-11-043-806-456
10	55	93.2	685	6	US-11-043-806-452
11	55	93.2	804	6	US-11-043-806-453
12	55	93.2	828	6	US-11-043-806-455
13	55	93.2	831	3	US-09-939-853A-97
14	55	93.2	831	3	US-09-939-853A-98
15	55	93.2	855	6	US-11-043-806-461
16	55	93.2	1000	6	US-11-043-806-457
17	55	93.2	1105	6	US-11-043-806-458
18	55	93.2	1150	4	US-10-296-733-1
19	55	93.2	1152	3	US-09-919-603-1
20	55	93.2	1169	5	US-10-317-821B-7
21	55	93.2	1170	4	US-10-020-141-12
22	55	93.2	1170	4	US-10-017-721-2
23	55	93.2	1170	4	US-10-021-660-114
24	55	93.2	1170	4	US-10-008-093-2
25	55	93.2	1170	4	US-10-295-027-1170
26	55	93.2	1170	4	US-10-211-462-38
27	55	93.2	1170	4	US-10-231-956A-482

28	55	93.2	1170	4	US-10-419-462-38	Sequence 38, Appl
29	55	93.2	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	55	93.2	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	55	93.2	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	55	93.2	1170	5	US-10-782-968-38	Sequence 38, Appl
33	55	93.2	1170	5	US-10-849-989-44	Sequence 44, Appl
34	55	93.2	1170	5	US-10-631-467-548	Sequence 548, App
35	55	93.2	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	55	93.2	1170	5	US-10-831-997-2	Sequence 2, Appl
37	55	93.2	1170	5	US-10-995-561-594	Sequence 594, App
38	55	93.2	1170	5	US-10-995-561-595	Sequence 595, App
39	55	93.2	1170	5	US-10-995-561-596	Sequence 596, App
40	55	93.2	1170	6	US-11-037-713-51	Sequence 51, Appl
41	55	93.2	1170	6	US-11-046-644-28	Sequence 28, Appl
42	55	93.2	1170	6	US-11-046-456-28	Sequence 28, Appl
43	45	76.3	15	4	US-10-285-394-153	Sequence 153, App
44	39	66.1	226	5	US-10-467-657-2428	Sequence 2428, Ap
45	39	66.1	498	4	US-10-425-114-49331	Sequence 49331, A

ALIGNMENTS

RESULT 1
US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Kruttsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/10535
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 93.2%; Score 55; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
| | | | | | | | | |
DB 1 FQGVQLQNVRFV 11

RESULT 2
US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J
; APPLICANT: Williams, Kevin J
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof in Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 40

; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match 93.2%; Score 55; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.032; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQGVQLQNVRFV 11
Db 190 FQGVQLQNVRFV 200
|||||

RESULT 3

US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match 93.2%; Score 55; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.032; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQGVQLQNVRFV 11
Db 190 FQGVQLQNVRFV 200
|||||

RESULT 4

US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 93.2%; Score 55; DB 5; Length 432;

Best Local Similarity 100.0%; Pred. No. 0.06; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218
|||||

RESULT 5

US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 93.2%; Score 55; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.06; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218
|||||

RESULT 6

US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match 93.2%; Score 55; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.064; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218
|||||

RESULT 7

US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 93.2%; Score 55; DB 3; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.065; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
| | | | | | | | | |
DB 261 FQGVQLQNVRFV 271

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 93.2%; Score 55; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 0.079; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
| | | | | | | | | |
DB 208 FQGVQLQNVRFV 218

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 93.2%; Score 55; DB 6; Length 578;
Best Local Similarity 100.0%; Pred. No. 0.082; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
| | | | | | | | | |
DB 208 FQGVQLQNVRFV 218

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 93.2%; Score 55; DB 6; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.099; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
| | | | | | | | | |
DB 208 FQGVQLQNVRFV 218

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 93.2%; Score 55; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
| | | | | | | | | |
DB 208 FQGVQLQNVRFV 218

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match 93.2%; Score 55; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218

RESULT 13

US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1

; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match 93.2%; Score 55; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218

RESULT 14

US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1

; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus

US-09-939-853A-98

Query Match 93.2%; Score 55; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218

RESULT 15

US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1

; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match 93.2%; Score 55; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
Db 208 FQGVQLQNVRFV 218

Search completed: June 6, 2006, 00:00:10
Job time : 78.6207 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-28

Perfect score: 59

Sequence: 1 FQGVLRVRFVA 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	57.6	132	US-10-953-349-39312	Sequence 39312, A
2	34	57.6	298	US-10-953-349-9347	Sequence 9347, Ap
3	34	57.6	331	US-10-953-349-34674	Sequence 34674, A
4	34	57.6	358	US-10-953-349-34673	Sequence 34673, A
5	34	57.6	382	US-10-953-349-34672	Sequence 34672, A
6	33	55.9	750	US-10-511-937-2413	Sequence 2413, Ap
7	33	55.9	1333	US-10-511-937-2992	Sequence 2992, Ap
8	32	54.2	186	US-10-953-349-27518	Sequence 27518, A
9	32	54.2	188	US-10-953-349-27517	Sequence 27517, A
10	32	54.2	213	US-10-953-349-27516	Sequence 27516, A
11	32	54.2	843	US-10-953-349-16527	Sequence 16527, A
12	32	54.2	962	US-10-953-349-16526	Sequence 16526, A
13	32	54.2	1014	US-10-953-349-16525	Sequence 16525, A
14	31	52.5	301	US-10-953-349-19576	Sequence 19576, A
15	31	52.5	320	US-10-953-349-19575	Sequence 19575, A
16	31	52.5	387	US-10-953-349-19574	Sequence 19574, A
17	31	52.5	438	US-10-953-349-28462	Sequence 28462, A
18	31	52.5	442	US-10-953-349-28461	Sequence 28461, A
19	31	52.5	453	US-10-953-349-28460	Sequence 28460, A
20	31	52.5	604	US-11-293-697-2816	Sequence 2816, Ap
21	31	52.5	608	US-11-203-828-5	Sequence 5, Appl
22	31	52.5	855	US-11-247-437-2	Sequence 2, Appl
23	30	50.8	178	US-11-293-697-3713	Sequence 3713, Ap
24	30	50.8	227	US-10-953-349-39138	Sequence 39138, A
25	30	50.8	251	US-10-953-349-3586	Sequence 3586, Ap

26	30	50.8	262	6	US-10-953-349-2388	Sequence 2388, Ap
27	30	50.8	265	6	US-10-953-349-39137	Sequence 39137, A
28	30	50.8	276	6	US-10-953-349-39136	Sequence 39136, A
29	30	50.8	277	6	US-10-953-349-3585	Sequence 3585, Ap
30	30	50.8	296	6	US-10-953-349-2387	Sequence 2387, Ap
31	30	50.8	329	6	US-10-953-349-24690	Sequence 24690, A
32	30	50.8	332	6	US-10-953-349-24689	Sequence 24689, A
33	30	50.8	404	6	US-10-953-349-24688	Sequence 24688, A
34	30	50.8	407	6	US-10-953-349-26200	Sequence 26200, A
35	30	50.8	724	7	US-11-293-697-3263	Sequence 3263, Ap
36	29	49.2	96	6	US-10-953-349-16838	Sequence 16838, A
37	29	49.2	139	7	US-11-293-697-3323	Sequence 3323, A
38	29	49.2	230	6	US-10-953-349-28990	Sequence 28990, A
39	29	49.2	234	6	US-10-953-349-23952	Sequence 22952, A
40	29	49.2	252	6	US-10-953-349-28989	Sequence 28989, A
41	29	49.2	282	6	US-10-953-349-22951	Sequence 22951, A
42	29	49.2	287	6	US-10-953-349-1067	Sequence 1067, Ap
43	29	49.2	307	6	US-10-953-349-28988	Sequence 28988, A
44	29	49.2	340	6	US-10-953-349-1066	Sequence 1066, Ap
45	29	49.2	437	6	US-10-504-120-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1

US-10-953-349-39312
; Sequence 39312, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39312
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-39312

Query Match 57.6%; Score 34; DB 6; Length 132;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy	1	FQGVLRVRFVA	12
Db	66	FEGLNNNVFIA	77

RESULT 2

US-10-953-349-9347
; Sequence 9347, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9347

Query Match 57.6%; Score 34; DB 6; Length 298;

```
Best Local Similarity 36.4%; Pred. No. 12;
Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQVRFA 12
Db 87 EGIQGVRFIS 97

RESULT 3
US-10-953-349-34674
; Sequence 34674, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match 57.6%; Score 34; DB 6; Length 331;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQVRFA 12
Db 225 QGVLFNIQVVS 235

RESULT 4
US-10-953-349-34673
; Sequence 34673, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

Query Match 57.6%; Score 34; DB 6; Length 358;
Best Local Similarity 54.5%; Pred. No. 14;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQVRFA 12
Db 252 QGVLFNIQVVS 262

RESULT 5
US-10-953-349-34672
; Sequence 34672, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
```

```
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34672
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34672

Query Match 57.6%; Score 34; DB 6; Length 382;
Best Local Similarity 54.5%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQVRFA 12
Db 276 QGVLFNIQVVS 286

RESULT 6
US-10-511-937-2413
; Sequence 2413, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2413
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2413

Query Match 55.9%; Score 33; DB 6; Length 750;
Best Local Similarity 40.0%; Pred. No. 51;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQVRF 10
Db 508 FENIQLNKF 517

RESULT 7
US-10-511-937-2992
; Sequence 2992, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
```


; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2992
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2992

Query Match 55.9%; Score 33; DB 6; Length 1333;
Best Local Similarity 50.0%; Pred. No. 96;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:||||:|:
Db 327 FRGVLEQLRWFA 338

RESULT 8

US-10-953-349-27518
; Sequence 27518, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27518
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-27518

Query Match 54.2%; Score 32; DB 6; Length 186;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:||||
Db 63 FSAALAEVRFVA 74

RESULT 9

US-10-953-349-27517
; Sequence 27517, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27517
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Triticum aestivum

US-10-953-349-27517

Query Match 54.2%; Score 32; DB 6; Length 188;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:||||
Db 65 FSAALAEVRFVA 76

RESULT 10

US-10-953-349-27516
; Sequence 27516, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27516
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-27516

Query Match 54.2%; Score 32; DB 6; Length 213;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:||||
Db 90 FSAALAEVRFVA 101

RESULT 11

US-10-953-349-16527
; Sequence 16527, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16527
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16527

Query Match 54.2%; Score 32; DB 6; Length 843;
Best Local Similarity 54.5%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
|:|:|:|:
Db 780 FKGLDNYVVF 790

RESULT 12

US-10-953-349-16526
; Sequence 16526, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16526
; LENGTH: 962
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16526

Query Match 54.2%; Score 32; DB 6; Length 962;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
|:|:|:|:|
DB 899 FKGILDNVYFV 909

RESULT 13

US-10-953-349-16525
; Sequence 16525, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16525
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16525

Query Match 54.2%; Score 32; DB 6; Length 1014;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
|:|:|:|:|
DB 951 FKGILDNVYFV 961

RESULT 14

US-10-953-349-19576
; Sequence 19576, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19576
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-19576

Query Match 52.5%; Score 31; DB 6; Length 301;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:|:|:|:|
DB 211 FQGLLRTLKQVA 222

RESULT 15

US-10-953-349-19575
; Sequence 19575, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19575
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-19575

Query Match 52.5%; Score 31; DB 6; Length 320;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFVA 12
|:|:|:|:|
DB 230 FQGLLRTLKQVA 241

Search completed: June 6, 2006, 00:12:56
Job time : 3.82414 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 91.1379 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-29

Perfect score: 60

Sequence: 1 FQGLANVRVVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 8:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*
10: Geneseq2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	12	4	AAB35362
2	55	91.7	12	4	AAB35352
3	55	91.7	12	4	AAB35378
4	55	91.7	12	6	ABG72834
5	55	91.7	240	8	ADL70641
6	55	91.7	432	8	ADQ39359
7	55	91.7	432	8	ADQ39357
8	55	91.7	459	4	AAU02916
9	55	91.7	466	3	AAU43602
10	55	91.7	546	4	AAU02915
11	55	91.7	548	7	ADN02474
12	55	91.7	555	4	AAU02914
13	55	91.7	731	3	AAU02913
14	55	91.7	1152	3	AAU00042
15	55	91.7	1152	5	AAU74771
16	55	91.7	1152	5	ABB82285
17	55	91.7	1170	4	AAU74450
18	55	91.7	1170	4	AAU90800
19	55	91.7	1170	5	AAE25030
20	55	91.7	1170	5	AAU75315
21	55	91.7	1170	6	ABP96780
22	55	91.7	1170	6	ABU03474
23	55	91.7	1170	6	ABG74673

24	55	91.7	1170	6	AAB36228	Aae36228 Human THB
25	55	91.7	1170	7	ABR62059	ABR62059 Human THB
26	55	91.7	1170	7	ADN39852	ADN39852 Cancer/an
27	55	91.7	1170	8	ADJ76124	ADJ76124 Marker ge
28	55	91.7	1170	8	ADJ75296	ADJ75296 Marker ge
29	55	91.7	1170	8	ADL70639	ADL70639 Human thr
30	55	91.7	1170	8	ADL70639	ADL70639 Human thr
31	55	91.7	1170	8	ADQ26070	ADQ26070 Thrombosop
32	55	91.7	1170	8	ADP54179	ADP54179 Human PRO
33	55	91.7	1170	8	ADQ39358	ADQ39358 Human myo
34	55	91.7	1170	8	ADQ39356	ADQ39356 Human myo
35	55	91.7	1170	8	ADQ39355	ADQ39355 Human myo
36	55	91.7	1170	9	ADZ21688	ADZ21688 Thrombosop
37	55	91.7	1170	9	AE887781	AE887781 Human thr
38	55	91.7	1170	9	AE846751	AE846751 Human thr
39	54	90.0	12	4	AAB35371	AAB35371 Alpha3bet
40	52	86.7	12	4	AAB35373	AAB35373 Alpha3bet
41	52	86.7	12	4	AAB35381	AAB35381 Alpha3bet
42	51	85.0	12	4	AAB35364	AAB35364 Alpha3bet
43	51	85.0	12	4	AAB35374	AAB35374 Alpha3bet
44	50	83.3	12	4	AAB35368	AAB35368 Alpha3bet
45	50	83.3	12	4	AAB35376	AAB35376 Alpha3bet

ALIGNMENTS

RESULT 1
AAB35362
ID AAB35362 standard; peptide; 12 AA.
AC AAB35362;
XX
DT 08-MAY-2001 (first entry)
DE Alpha3bet integrin binding peptide #27.
XX
XX Alpha3bet integrin; angiogenesis; cell proliferation; cancer;
KW Alpha3bet integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
XX of the invention
XX
XX Sequence 12 AA;

```

Query Match      100.0%; Score 60; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVLANVRVF 12
       |||||
Db      1 FQGVLANVRVF 12

RESULT 2
AAB35352
ID AAB35352 standard; peptide; 12 AA.
XX
AC AAB35352;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3betal integrin binding peptide #17.
XX
KW Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;
XX
DR WPI; 2001-182656/18.
XX
PT New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
PS Claim 4; Page 34; 84pp; English.
XX
CC The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
SQ Sequence 12 AA;

Query Match      91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLANVRVF 12
       |||||
Db      1 FQGVLANVRVF 12

RESULT 4
AAB35378
ID AAB35378 standard; peptide; 12 AA.
XX
AC AAB35378;
XX
DT 24-FEB-2003 (first entry)
XX
DE Thrombospondin-1 sequence containing synthetic peptide.
XX
KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
OS Synthetic.
XX
PN WO200281630-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010535.
XX
PR 06-APR-2001; 2001US-0281994P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roberts DD, Kruttsch HC;

```

XX WPI; 2003-103329/09.
 XX
 PT A new diagnosis for cancer other than prostate cancer in a mammal useful
 PT to detect cancer including lung cancer, particularly small cell lung
 PT cancer and melanoma comprises detecting semenogelin in a sample.
 XX
 PS Example 1; Page 14; 32pp; English.
 XX
 CC The invention relates to diagnosing cancer other than prostate cancer in
 CC a male mammal, comprising assaying a test sample for increased level of
 CC semenogelin, or cancer in a female by assaying for the presence of
 CC semenogelin. Administering a semenogelin protein or polypeptide fragment
 CC or a semenogelin-specific antibody or active fragment, or a recombinant
 CC vector expressing the protein or antibody, is useful for inducing an
 CC immune response to a cancer in a mammal, where the cancer is not prostate
 CC cancer and semenogelin is a marker. The invention is used to diagnose
 CC cancer, particularly of epithelial origin such as lung cancer, papillary
 CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
 CC (SCLC), or a melanoma. The present sequence represents the amino acid
 CC sequence of the thrombospondin-1 sequence containing synthetic peptide
 CC which binds to alpha-3-beta-1 integrin
 XX
 SQ Sequence 12 AA;
 Query Match 91.7%; Score 55; DB 6; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.00063;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVLANVRVF 12
 DB 1 FQGVLANVRVF 12
 RESULT 5
 ID ADL70641 standard; protein; 240 AA.
 XX
 AC ADL70641;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human thrombospondin-1 N-terminal domain.
 XX
 KW Human; thrombospondin-1; epitope; cancer; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 23..32
 FT /note= "Heparin binding region"
 FT Region 77..82
 FT /note= "Heparin binding region"
 FT Region 151..164
 FT /note= "Fibrinogen binding region"
 XX
 PN WO2004018995-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 20-AUG-2003; 2003WO-US026023.
 XX
 PR 23-AUG-2002; 2002US-0405494P.
 PR 21-APR-2003; 2003US-00419462.
 XX
 PA (WILL/) WILLIAMS K J.
 XX
 PI Williams KJ;
 XX
 DR WPI; 2004-226901/21.
 XX
 XX New purified thrombospondin fragment extracted from a body fluid, useful
 PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma

or leukemia or as calibrators, indicators, immunogens and analytes.
 Disclosure; SEQ ID NO 40; 76pp; English.
 The present sequence is that of the N-terminal domain of human
 thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
 (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
 clinical assays for cancer and for generation of antibodies and other
 binding agents. A method that distinguishes TSP from a TSP fragment or
 portion as a target for a binding molecule, e.g. an antibody, to
 obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
 epitope present in TSP but not in the fragment or portion to obtain a
 quantitation of TSP only; and (3) using the difference between (1) and
 (2) as a quantitation of the amount of TSP fragment or portion. Suitable
 epitopes are provided ADL70602-ADL70638. Detection or quantification of
 the TSP fragment or portion is performed in order to detect the presence,
 or monitor the course, of a disease or condition selected from cancer,
 renal failure, renal disease, atopic dermatitis, vasculitis, acute
 vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
 infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
 nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
 syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
 rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
 thrombocytopenic purpura, purpura, an inflammatory condition, a condition
 associated with clotting, a condition associated with platelet
 activation, a condition associated with intravascular platelet
 activation, a condition associated with consumption of platelets, heparin
 -induced thrombocytopenia, disseminated intravascular coagulation,
 intravascular coagulation, extravascular coagulation, a condition
 associated with endothelial activation, a condition associated with
 production and/or release of thrombospondin and/or a thrombospondin
 fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 reaction, an aspartame reaction, atopic dermatitis, eczema,
 hypersensitivity, scleroderma, conditions associated with plugging of
 vessels, a condition associated with a cryofibrinogen, a condition
 associated with a cryoglobulin, and a condition associated with an anti-
 cardiolipin antibody. The cancer is selected from adenoma,
 adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 the respiratory system, circulatory system, musculoskeletal system,
 muscle, bone, a joint, tendon or ligament, digestive system, liver or
 biliary system, pancreas, head, neck, endocrine system, reproductive
 system (male or female), genitourinary system, kidney, urinary tract,
 sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 well-differentiated cancer or a moderately differentiated cancer.

Query Match 91.7%; Score 55; DB 8; Length 240;
 Best Local Similarity 91.7%; Pred. No. 0.018;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FQGVLANVRVF 12
 DB 190 FQGVLANVRVF 201
 RESULT 6
 ADQ39359
 ID ADQ39359 standard; protein; 432 AA.
 XX
 AC ADQ39359;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human myocardial infarction-associated gene derived protein, SEQ ID 1022.
 XX
 KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 KW cardiac; gene therapy; human.


```

ID AAU02916 standard; protein; 459 AA.
AC AAU02916;
XX
XX
XX 12-SEP-2001 (first entry)
XX
XX Angiotensin converting enzyme (ACEV) splice variant protein #16.
XX
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX platelet-derived endothelial cell growth factor; cardiovascular disease;
XX cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX myocardial infarction; coronary arterial thrombosis; renal disease;
XX diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
XX multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
XX vascular disorder; asbestosis.
XX
XX Homo sapiens.
XX
XX WO200136632-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-IL000766.
XX
XX 17-NOV-1999; 99IL-00132978.
XX 10-DEC-1999; 99IL-00133455.
XX (COMP-) COMPUGEN LTD.
XX
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX N-PSDB; AAS06016.
XX
XX Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
XX capable of binding to the variant and to detect anti-variant antibodies.
XX
XX Claim 4; Fig 16; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
XX polypeptide receptor 2. The polypeptides and their associated nucleic
XX acids are useful for identification of variant sequences and detection of
XX candidate compounds capable of binding the molecules. The sequences of
XX the invention can be used in the treatment and diagnosis of various
XX disorders including cardiovascular diseases such as arteriosclerosis,
XX myocardial infarction and coronary arterial thrombosis, renal diseases
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX immune disorders such as immune complex nephritis, multiple sclerosis,
XX cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
XX as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
XX
XX Sequence 459 AA;
XX
XX Query Match 91.7%; Score 55; DB 4; Length 459;
XX Best Local Similarity 91.7%; Pred. No. 0.036;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 FQGVLANVRVF 12
XX ||||| |||||
XX 208 FQGVLANVRVF 219
XX
XX Db
XX
XX RESULT 9
XX AAB43602

```

```

ID AAB43602 standard; protein; 466 AA.
XX
XX AAB43602;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated protein sequence SEQ ID NO:1047.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
XX antiidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX vasotropic; anipsoriatic; antiangiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening.
XX
XX Homo sapiens.
XX
XX WO200055350-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005882.
XX
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX N-PSDB; AAC77811.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX
XX Claim 11; Page 1636-1638; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnery; immunomodulator;
XX antiidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
XX nootropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 466 AA;
XX
XX Query Match 91.7%; Score 55; DB 3; Length 466;
XX Best Local Similarity 91.7%; Pred. No. 0.037;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 FQGVLANVRVF 12
XX ||||| |||||
XX 261 FQGVLANVRVF 272
XX
XX Db
XX
XX RESULT 9
XX AAB43602

```

```

RESULT 10
AAU02915
ID AAU02915 standard; protein; 546 AA.
XX
AC AAU02915;
XX
DT 12-SEP-2001 (first entry)
XX
DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonarctoidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX
OS Homo sapiens.
XX
PN WO200136632-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-1L000766.
XX
PR 17-NOV-1999; 99IL-00132978.
PR 10-DEC-1999; 99IL-00133455.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX
XX WPI; 2001-336004/35.
XX
XX N-PSDB; AAS06015.
XX
XX Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
XX capable of binding to the variant and to detect anti-variant antibodies.
XX
XX Claim 4; Fig 15; 51pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
XX polypeptide receptor 2. The polypeptides and their associated nucleic
XX acids are useful for identification of variant sequences and detection of
XX candidate compounds capable of binding the molecules. The sequences of
XX the invention can be used in the treatment and diagnosis of various
XX disorders including cardiovascular diseases such as arteriosclerosis,
XX myocardial infarction and coronary arterial thrombosis, renal diseases
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX immune disorders such as immune complex nephritis, multiple sclerosis,
XX cancer, sarcoidosis, nonarctoidotic pulmonary granulomatous diseases such
XX as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
XX
XX Sequence 546 AA;
XX
XX Query Match 91.7%; Score 55; DB 4; Length 546;
XX Best Local Similarity 91.7%; Pred. No. 0.044;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLANVRVVF 12
XX ||||| |||||
XX Db 208 FQGVLANVRVVF 219
XX
XX RESULT 12
XX AAU02914
XX ID AAU02914 standard; protein; 555 AA.
XX
XX AC AAU02914;
XX
XX DT 12-SEP-2001 (first entry)
XX
XX DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX platelet-derived endothelial cell growth factor; cardiovascular disease;
XX cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX myocardial infarction; coronary arterial thrombosis; renal disease;
XX diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
XX multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX nonarctoidotic pulmonary granulomatous disease; endothelial abnormality;
XX
XX Sequence 548 AA;
XX
XX Query Match 91.7%; Score 55; DB 7; Length 548;
XX Best Local Similarity 91.7%; Pred. No. 0.044; 1; Indels 0; Gaps 0;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLANVRVVF 12
XX ||||| |||||
XX Db 208 FQGVLANVRVVF 219
XX
XX RESULT 11
XX ADN02474
XX ID ADN02474 standard; protein; 548 AA.
XX
XX AC ADN02474;
XX
XX DT 17-JUN-2004 (first entry)
XX
XX DE TSF polypeptide.
XX
XX KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX
XX OS Homo sapiens.
XX
XX PN CN1401387-A.
XX
XX PD 12-MAR-2003.
XX
XX PF 21-AUG-2002; 2002CN-00129408.
XX
XX PR 21-AUG-2002; 2002CN-00129408.
XX
XX (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
XX
XX Han Z, Liu P;
XX
XX WPI; 2003-469302/45.
XX
XX N-PSDB; ADN02475.
XX
XX Tumor suppressing polypeptide TSF and gene therapy vector composition.
XX
XX Claim 2; SEQ ID NO 1; 13pp; Chinese.
XX
XX The present invention relates to a novel recombinant adenovirus vector
XX mediated anti-neoplastic composition is prepared through cloning the cDNA
XX sequence from the human peripheral blood cell by specific primer and
XX reverse transcription-polymerase chain reaction (RT-PCR) method for
XX coding TSF polypeptide, construction in human embryonic kidney cell 293
XX by AdEasy system, and packaging and expressing the recombinant adenovirus
XX vector of TSF. It can suppress the growth and transfer of cancer. The
XX present sequence represents the TSF polypeptide.
XX
XX Sequence 548 AA;
XX
XX Query Match 91.7%; Score 55; DB 7; Length 548;
XX Best Local Similarity 91.7%; Pred. No. 0.044; 1; Indels 0; Gaps 0;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLANVRVVF 12
XX ||||| |||||
XX Db 208 FQGVLANVRVVF 219
XX
XX RESULT 12
XX AAU02914
XX ID AAU02914 standard; protein; 555 AA.
XX
XX AC AAU02914;
XX
XX DT 12-SEP-2001 (first entry)
XX
XX DE Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX
XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX platelet-derived endothelial cell growth factor; cardiovascular disease;
XX cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX myocardial infarction; coronary arterial thrombosis; renal disease;
XX diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
XX multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX nonarctoidotic pulmonary granulomatous disease; endothelial abnormality;
XX
XX Sequence 548 AA;
XX
XX Query Match 91.7%; Score 55; DB 7; Length 548;
XX Best Local Similarity 91.7%; Pred. No. 0.044; 1; Indels 0; Gaps 0;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FQGVLANVRVVF 12
XX ||||| |||||
XX Db 208 FQGVLANVRVVF 219
XX

```


KW vascular disorder; asbestosis.

XX Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

XX 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

XX N-PSDB; AAS06014.

XX Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 4; Fig 14; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis

XX Sequence 555 AA;

Query Match 91.7%; Score 55; DB 4; Length 555;

Best Local Similarity 91.7%; Pred. NO. 0.045;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12

DB 208 FQGVLANVRVVF 219

RESULT 13

AAU02913

ID AAU02913 standard; protein; 731 AA.

XX AAU02913;

DT 12-SEP-2001 (first entry)

DE Angiotensin converting enzyme (ACEV) splice variant protein #13.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

KW vascular disorder; asbestosis.

XX Homo sapiens.

XX WO200136632-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-IL000766.

XX 17-NOV-1999; 99IL-00132978.

XX 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUGEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

XX N-PSDB; AAS06013.

XX Novel alternative splicing variants e.g. variant of angiotensin converting enzyme (ACEV), useful in identifying candidate compounds capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 4; Fig 13; 519pp; English.

XX The sequence represents an angiotensin converting enzyme splice variant (ACEV) polypeptide. The polypeptides of the invention include variants of granulocyte colony stimulating factor receptor, glucagon, interleukin 6, platelet-derived endothelial cell growth factor, cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal polypeptide receptor 2. The polypeptides and their associated nucleic acids are useful for identification of variant sequences and detection of candidate compounds capable of binding the molecules. The sequences of the invention can be used in the treatment and diagnosis of various disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis

XX Sequence 731 AA;

Query Match 91.7%; Score 55; DB 4; Length 731;

Best Local Similarity 91.7%; Pred. NO. 0.061;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12

DB 208 FQGVLANVRVVF 219

RESULT 14

AAB00042

ID AAB00042 standard; protein; 1152 AA.

XX AAB00042;

DT 08-NOV-2000 (first entry)

DE Human thrombospondon-1 (TSP-1).

XX TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
KW diabetic retinopathy; corneal graft rejection; glaucoma.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 361..416

XX /label= Type 1 repeat region

CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases (e.g.
CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasias,
CC rebeosis), Osler-Webber syndrome, myocardial angiogenesis,
CC telangiectasia, plaque neovascularisation, haemophilic joints,
CC angiofibroma or wound granulation. The composition induces apoptosis and
CC inhibits neovascularisation in the tumour cells. This amino acid sequence
CC represents human thrombospondin-1 (TSP-1), on which the recombinant
XX proteins of the invention are based

SQ Sequence 1152 AA;

Query Match 91.7%; Score 55; DB 5; Length 1152;
Best Local Similarity 91.7%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 12
| | | | | | | | | |
Db 190 FQGVLANVRVF 201

Search completed: June 5, 2006, 22:24:59
Job time : 91.1379 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 13.9655 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-29
Perfect score: 60
Sequence: 1 FQGVLANVRVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	55	91.7	229	S57957	thrombospondin 1 -
2	55	91.7	1170	TSHUP1	thrombospondin 1 p
3	55	91.7	1170	A40558	thrombospondin 1 p
4	40	66.7	304	T34271	hypothetical prote
5	38	63.3	1054	H69377	reverse gyrase (to
6	37	61.7	400	T24890	hypothetical prote
7	37	61.7	467	D84938	H+-transporting tw
8	37	61.7	628	AF1108	transporter homolo
9	37	61.7	628	AG1469	transporter homolo
10	37	61.7	1172	TSHUP2	thrombospondin 2 p
11	37	61.7	1172	A42587	thrombospondin 2 p
12	37	61.7	1285	T14171	ataxin-2 - mouse
13	36	60.0	248	F69113	hypothetical prote
14	36	60.0	425	D88115	protein F53C3.11 f
15	36	60.0	481	T15657	hypothetical prote
16	36	60.0	1839	S77626	mammalian C-5-epi
17	35	58.3	111	B71115	hypothetical prote
18	35	58.3	152	A24740	C02C2.2 protein -
19	35	58.3	224	A64599	hydrogenase (SC 1.
20	35	58.3	224	D71915	hydrogenase, cyto
21	35	58.3	298	C87403	fdhd protein (limp
22	35	58.3	385	S56224	hypothetical prote
23	35	58.3	399	AG2221	hypothetical prote
24	35	58.3	436	D64056	branched-chain ami
25	35	58.3	563	T36704	probable dehydroge
26	35	58.3	608	AG1259	GTP-binding protei
27	35	58.3	608	A11621	GTP-binding protei
28	35	58.3	1178	A39804	thrombospondin pre
29	35	58.3	1288	T42756	5-oxoprolinase (AT

30	58.3	1829	2	AE1864	hypothetical prote
31	56.7	111	2	S52596	probable membrane
32	56.7	206	2	T45249	probable amidotran
33	56.7	247	1	A64590	probable 3-oxoacyl
34	56.7	247	2	B71923	3-oxoacyl-l-acyl-ca
35	56.7	268	2	B87397	hypothetical prote
36	56.7	343	2	JN0249	vancomycin resist
37	56.7	415	2	G90705	probable transport
38	56.7	415	2	B85556	probable transport
39	56.7	415	2	G64790	ybdG protein - Esc
40	56.7	467	2	A85363	probable calmoduli
41	56.7	469	2	T33595	hypothetical prote
42	56.7	552	2	AC0709	probable ptrBa pro
43	56.7	683	2	AC0741	oligopeptidase B (
44	56.7	686	1	E64946	oligopeptidase B (
45	56.7	686	2	C90948	proteinase II (imp

ALIGNMENTS

RESULT 1

S57957
thrombospondin 1 - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S57957
R:Lafuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.
submitted to the EMBL Data Library, July 1995
A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expres
A:Reference number: S57955
A:Accession: S57957
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <LAF>
A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:G899228; PIR:
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v

Query Match 91.7%; Score 55; DB 2; Length 229;
Best Local Similarity 91.7%; Pred. No. 0.0033;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 12
DB 190 FQGVLANVRVF 201

RESULT 2

TSHUP1
thrombospondin 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: A26155; A34274; A30140; A25812; A05172; A42927
R:Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple
A:Reference number: A26155; MUID:87057617; PMID:2430973
A:Accession: A26155
A:Molecule type: mRNA
A:Residues: 1-1170 <LAW>
A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:G37137; PIDN:
A>Note: parts of this sequence, including the amino end of the mature protein, were det
R:laherty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA
A:Reference number: A34274; MUID:89291870; PMID:2544587
A:Accession: A34274
A:Molecule type: DNA
A:Residues: 1-166 <LAH>
A:Cross-references: UNIPARC:UPI000017428F; GB:J04835
R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,
J. Cell Biol. 108, 729-736, 1989
A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in t


```
A;Gene: CBSP:P46C8.8
A;Introns: 24/1; 56/1; 101/1; 123/1; 149/3; 219/1; 275/1

Query Match      66.7%; Score 40; DB 2; Length 304;
Best Local Similarity 70.0%; Pred. No. 4.1;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 10
    |||:||||:|
Db 22 FQGVLANAKF 31

RESULT 5
H69377
reverse gyrase (top-RG) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69377
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Uutterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: H69377
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1054 <KLE>
A;Cross-references: UNIPROT:Q29238; UNIPARC:UPI0000056EB5; GB:AE001033; GB:AE000782; NID
F;78-85/Region: nucleotide binding; P-loop
F;257-262/Region: nucleotide-binding motif A (P-loop)
F;261-263/Region: DEAD motif #status atypical

Query Match      63.3%; Score 38; DB 2; Length 1054;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 11
    |||:||||:|
Db 565 FHGVLVNGRFV 575

RESULT 6
T24890
hypothetical protein Tl3F3.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24890
R;Gardner, A.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19948
A;Accession: T24890
A;Status: preliminary; translated from GB/EMBL/DDBU
A;Molecule type: DNA
A;Residues: 1-400 <WIL>
A;Cross-references: UNIPROT:O18086; UNIPARC:UPI0000080DD5; EMBL:Z93389; PIDN:CA507670.1;
C;Genetics:
A;Experimental source: clone Tl3F3
A;Gene: CBSP:Tl3F3.2
A;Map position: 5
A;Introns: 50/2; 87/1; 141/3; 171/3; 242/3; 315/1; 360/3

Query Match      61.7%; Score 37; DB 2; Length 400;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 12
    ||:||||:|
Db 129 FKSLLAEVRYIF 140
```

```
RESULT 7
D84938
H+-transporting two-sector ATPase (EC 3.6.3.14), flagellum-specific [imported] - Buchne
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C;Accession: D84938
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
A;Reference number: A84930; MUID:20445173; PMID:10993077
A;Accession: D84938
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <STO>
A;Cross-references: UNIPARC:UPI000005E44F; GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: fliI; BU076
C;Superfamily: H(+)-transporting ATP synthase; H+-transporting ATP synthase alpha chain
C;Keywords: Hydrolase

Query Match      61.7%; Score 37; DB 2; Length 467;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 10
    ||:||||:|
Db 10 FRGILNLRVF 19

RESULT 8
AF1108
transporter homolog lmo0269 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF1108
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A;Title: Comparative Genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1108
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-628 <GLA>
A;Cross-references: UNIPROT:Q8YA90; UNIPARC:UPI0000054FB6; GB:NC_003210; PIDN:CAD000796.
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0269

Query Match      61.7%; Score 37; DB 2; Length 628;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLNVRVFVF 12
    ||:||||:|
Db 580 GLIANAKFVF 589

RESULT 9
AG1469
transporter homolog lin0294 [imported] - Listeria innocua (strain Clipl1262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1469
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
```

A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria species*.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AG1469
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-628 <GLA>
 A:Cross-references: UNIPROT:Q92FI4; UNIPARC:UPI00000CC1A0; GB:AL592022; PIDN:CAC95527.1;
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: lin0294

Query Match 61.7%; Score 37; DB 2; Length 628;
 Best Local Similarity 60.0%; Pred. No. 34; Mismatches 1; Indels 0; Gaps 0;
 Matches 6; Conservative 3;

QY 3 GVLNVRVRFV 12

Db 580 GLIANAKVFF 589

RESULT 10

TSHUP2

thrombospondin 2 precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
 C:Accession: A47379; A42173
 R:LaBell, T.L.; Byers, P.H.
 Genomics 17, 225-229, 1993
 A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
 A:Reference number: A47379; MUID:94010892; PMID:8406456
 A:Accession: A47379
 A:Molecule type: mRNA
 A:Residues: 1-1172 <LAB>
 A:Cross-references: UNIPROT:P35442; UNIPARC:UPI000046680; GB:L12350; NID:g307505; PIDN:
 R:LaBell, T.L.; Milewicz, D.J.; Distèche, C.M.; Byers, P.H.
 Genomics 12, 421-429, 1992
 A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression c
 A:Reference number: A42173; MUID:92217961; PMID:1559694
 A:Accession: A42173
 A:Molecule type: mRNA
 A:Residues: 560-1172 <LA2>
 A:Cross-references: UNIPARC:UPI00001742C1; GB:M81339
 A:Experimental source: fibroblast
 A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
 C:Genetics:
 A:Gene: GDB:THBS2; TSP2
 A:Cross-references: GDB:128789; OMIM:188061
 A:Map position: 6q27-6q27
 C:Complex: homotrimer, disulfide linked
 C:Function:
 A:Description: participates in cell migration and adhesion, and in platelet aggregation
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
 C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
 F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F:553-588/Domain: EGF homology <EGF1>
 F:652-691/Domain: EGF homology <EGF>
 F:928-930/Region: cell attachment (R-G-D) motif
 F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predi
 F:167,226/Disulfide bonds: #status predicted
 F:266,270/Disulfide bonds: interchain #status predicted
 F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 61.7%; Score 37; DB 1; Length 1172;
 Best Local Similarity 58.3%; Pred. No. 64;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVLNVRVRFV 12
 Db 202 FRGLLQNVHLVF 213

RESULT 11

A42587

thrombospondin 2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A42587; A39851
 R:laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
 J. Biol. Chem. 267, 3274-3281, 1992
 A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
 A:Reference number: A42587; MUID:92147683; PMID:1371115
 A:Accession: A42587
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-1172 <LAH>
 A:Cross-references: UNIPROT:Q03350; UNIPARC:UPI000029847; GB:L07803; GB:M87275; NID:g34
 A:Note: sequence extracted from NCBI backbone (NCBIP:81502)
 R:Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
 J. Biol. Chem. 266, 12821-12824, 1991
 A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
 A:Reference number: A39851; MUID:91302287; PMID:1712771
 A:Accession: A39851
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-873 <BOR>
 A:Cross-references: UNIPARC:UPI000016D077; GB:M64866; NID:g201994; PIDN:AAA0432.1; PID:
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
 C:Keywords: calcium binding; glycoprotein
 F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F:553-588/Domain: EGF homology <EGF1>
 F:652-691/Domain: EGF homology <EGF>

Query Match 61.7%; Score 37; DB 2; Length 1172;
 Best Local Similarity 58.3%; Pred. No. 64;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVLNVRVRFV 12

Db 202 FRGLLQNVHLVF 213

RESULT 12

T14171

ataxin-2 - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T14171
 R:Nechiporuk, T.T.; Figueroa, K.; Sahba, S.; Nechiporuk, A.V.; Pulst, S.M.
 submitted to the EMBL Data Library, January 1998
 A:Description: Mouse homolog of the SCA2 gene.
 A:Reference number: Z17900
 A:Accession: T14171
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-1285 <NEC>
 A:Cross-references: UNIPROT:070305; UNIPARC:UPI0000027D7E; EMBL:AF041472; NID:g3005019;
 C:Genetics:
 A:Gene: SCA2

Query Match 61.7%; Score 37; DB 2; Length 1285;
 Best Local Similarity 63.6%; Pred. No. 70;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVLNVRVRFV 11

Db 229 FDGIYANVRMV 239

RESULT 13

F69113
hypothetical protein MTH1844 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: F69113
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A;Reference number: A69000; MUID:98037514; PMID:19371463
A;Accession: F69113
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-248 <MTH>
A;Cross-references: UNIPROT:O27872; UNIPARC:UPI0000062C0F; GB:AE000937; GB:AE000666; NID
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1844

Query Match 60.0%; Score 36; DB 2; Length 248;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANRVFV 11
|||:|:|:
DB 14 FQNLADIRFL 24

RESULT 14

D88115
protein F53C3.11 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D88115
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D88115
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-425 <STO>
A;Cross-references: UNIPROT:Q9TX77; UNIPARC:UPI0000078A71; GB:chr_II; PIDN:AC67451.1; PT
C;Genetics:
A;Gene: F53C3.11
A;Map position: 2

Query Match 60.0%; Score 36; DB 2; Length 425;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLANRVFV 12
|||:|:|:
DB 326 FENIIANVKRVF 337

RESULT 15

T15657
hypothetical protein C27D9.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15657
R;Chissoe, S.
submitted to the EMBL Data Library, February 1996
A;Description: The sequence of C. elegans cosmid C27D9.
A;Reference number: Z18384
A;Accession: T15657

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-481 <CHI>
A;Cross-references: UNIPROT:Q18255; UNIPARC:UPI000007D9E3; EMBL:U49955; NID:g1208838; P
C;Genetics:

A;Gene: CRSP:C27D9.1
A;Introns: 24/3; 65/3; 133/2; 196/2; 421/3
C;Superfamily: Caenorhabditis elegans hypothetical protein C27D9.1

Query Match 60.0%; Score 36; DB 2; Length 481;
Best Local Similarity 58.3%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLANRVFV 12
|||:|:|:
DB 347 FQNVCANVKRIF 358

Search completed: June 5, 2006, 22:45:03
Job time : 14.9655 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 108.931 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-29
Perfect score: 60
Sequence: 1 FQGVLANVRVVF 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 7.2.*
1: uniprot_sprot.*
2: uniprot_tmbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	91.7	229	Q28194_BOVIN	Q28194 bos taurus
2	55	91.7	496	Q7SY84_XENLA	Q7SY84 xenopus lae
3	55	91.7	1170	TSPI1_BOVIN	Q28178 bos taurus
4	55	91.7	1170	TSPI1_HUMAN	P07996 homo sapien
5	55	91.7	1170	TSPI1_MOUSE	P35441 mus musculus
6	55	91.7	1170	Q3T40_MOUSE	Q3T40 mus musculus
7	55	91.7	1170	Q71SA3_RAT	Q71SA3 rattus norv
8	55	91.7	1171	Q80YQ1_MOUSE	Q80YQ1 mus musculus
9	55	91.7	1171	Q8CGB2_MOUSE	Q8CGB2 mus musculus
10	55	91.7	1173	TSPI1_XENLA	P35448 xenopus lae
11	55	91.7	1225	Q59E99_HUMAN	Q59E99 homo sapien
12	50	83.3	1090	Q5SPG5_BRARE	Q5SPG5 brachydanio
13	50	83.3	1193	Q4S758_TETNG	Q4S758 tetraodon n
14	49	81.7	249	Q5U903_PIG	Q5U903 sus scrofa
15	49	81.7	1171	Q4RLR5_TETNG	Q4RLR5 tetraodon n
16	45	75.0	1034	Q4RQ74_TETNG	Q4RQ74 tetraodon n
17	44	73.3	727	Q8PRY3_METWA	Q8PRY3 methanosarc
18	44	73.3	730	Q467V1_METFA	Q467V1 methanosarc
19	44	73.3	733	Q8TLX6_METAC	Q8TLX6 methanosarc
20	43	71.7	1113	Q2LTC9_9DELT	Q2LTC9 syntrophus
21	42	70.0	713	Q3FIU8_9BURK	Q3FIU8 burkholderi
22	42	70.0	713	Q44XL2_9BURK	Q44XL2 burkholderi
23	42	70.0	713	Q4LLM8_9BURK	Q4LLM8 burkholderi
24	40	66.7	321	Q612H2_CAEBR	Q612H2 caenorhabdi
25	40	66.7	716	Q5Z2D3_NOCPA	Q5Z2D3 nocardia fa
26	40	66.7	1034	Q4DWH6_TRYCR	Q4DWH6 trypanosoma
27	40	66.7	1034	Q7Z2B9_TRYCR	Q7Z2B9 trypanosoma
28	39	65.0	929	Q4ILJ7_GIBZE	Q4ILJ7 gibberella
29	38	63.3	175	P90504_HVHR	P90504 human herpe
30	38	63.3	189	Q31IL6_THICR	Q31IL6 thiomicrosp
31	38	63.3	212	Q5FA21_NEIG1	Q5FA21 neisseria g

32	38	63.3	214	2	Q3QD05_9GAMM	Q3QD05 shewanella
33	38	63.3	216	2	Q3P5P7_9GAMM	Q3P5P7 shewanella
34	38	63.3	261	2	Q3C7L8_9CLOT	Q3C7L8 alkaliphilu
35	38	63.3	378	2	Q563V1_XENLA	Q563V1 xenopus lae
36	38	63.3	380	2	Q563S6_9PIPI	Q563S6 xenopus ruw
37	38	63.3	380	2	Q563S7_9PIPI	Q563S7 xenopus ruw
38	38	63.3	380	2	Q563S8_9PIPI	Q563S8 xenopus and
39	38	63.3	380	2	Q563S9_9PIPI	Q563S9 xenopus and
40	38	63.3	380	2	Q563T0_9PIPI	Q563T0 xenopus ruw
41	38	63.3	380	2	Q563T1_9PIPI	Q563T1 xenopus lon
42	38	63.3	380	2	Q563T2_9PIPI	Q563T2 xenopus lon
43	38	63.3	380	2	Q563T3_9PIPI	Q563T3 xenopus lon
44	38	63.3	380	2	Q563T4_9PIPI	Q563T4 xenopus ami
45	38	63.3	380	2	Q563T5_9PIPI	Q563T5 xenopus ami

ALIGNMENTS

RESULT 1
Q28194_BOVIN
ID Q28194_BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96311130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCP19>3.3.CO;2-O;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
Reige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
J. Cell. Physiol. 167:164-172(1996).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
EMBL: X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON_TER 1
FT NON_TER 229 229
SQ SEQUENCE 229 AA; 90D9EBCE4E6B669C CRC64;
Query Match 91.7%; Score 55; DB 2; Length 229;
Best Local Similarity 91.7%; Pred. No. 0.027;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGVLANVRVVF 12
||| |||||
Db 190 FQGVLANVRVVF 201
RESULT 2
Q7SY84_XENLA
ID Q7SY84_XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-FEB-2006, entry version 12.

DE MGC64438 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OX Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8355;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Whole;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeng B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Whole;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Whole;
RC Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; BC054970; AAH54970.1; -; mRNA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR01007; VWF_C.
DR Pfam; PF00090; TSP_1; 2.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS0092; TSP1; 2.
DR PROSITE; PS01208; VWF_C 1; UNKNOWN_1.
DR PROSITE; PS0184; VWF2; 1.
SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7EF51B CRC64;

Query Match 91.7%; Score 55; DB 2; Length 496;
Best Local Similarity 91.7%; Pred. No. 0.061; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQGLANVRVVF 12
||||| |||||

Db 214 FQGLQNVRFV 225
RESULT 3
ID TSP1_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-MAR-2006, entry version 56.
DE Thrombospondin-1 precursor.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RX [1]
RN NUCLEOTIDE SEQUENCE [MRNA].
RP STRAIN=Holstein; TISSUE=Tooth;
RC MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RX Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
RT odontoblasts and predentin.";
RL Biochim. Biophys. Acta 1382:17-22(1998).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
RP TISSUE=Aortic endothelium;
RC Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-1b/beta-3. May play a role in dentinogenesis
CC and/or maintenance of dentin and dental pulp.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- TISSUE SPECIFICITY: Odontoblasts.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWF domain.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; AB005287; BAA21115.1; -; mRNA.
DR EMBL; X87618; CAA60950.1; -; mRNA.
DR EMBL; X87619; CAA60951.1; -; mRNA.
DR PIR; S55501; S55501.
DR HSSP; P07996; 1LSL.
DR SMR; Q28178; 549-1169.
DR GlycoSuiteDB; Q28178; -.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF_3.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 13.

DR PFAM: PF05735; TSP_C; 1.
DR PFAM: PF00093; WVC; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; WVC; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00266; EGF_3; 2.
DR PROSITE: PS00092; TSP1; 3.
DR PROSITE: PS01208; WVC; 1; 1.
DR PROSITE: PS0184; WVC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT REGION 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT CARBOHYD 1085 1085
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT DISULFID 519 531
FT DISULFID 551 562
FT DISULFID 556 572
FT DISULFID 575 586
FT DISULFID 592 608
FT DISULFID 599 617
FT DISULFID 620 644
FT DISULFID 650 663
FT DISULFID 657 676
FT DISULFID 678 689
FT DISULFID 705 718
FT DISULFID 718 738
FT DISULFID 754 774
FT DISULFID 777 797
FT DISULFID 813 833
FT DISULFID 836 856
FT DISULFID 874 894
FT DISULFID 910 930
FT DISULFID 946 1167
FT CONFLICT 805 805
FT SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;
Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred.No. 0.15; Mismatches 11; Conservative 0; Gaps 0;
Matches 11; Indels 1; Indels 0; Gaps 0;
QY 1 FQGVLANVRVF 12
Db 208 FQGVLANVRVF 219
RESULT 4
TSP1_HUMAN
ID TSP1_HUMAN STANDARD; PRT; 1170 AA.
AC P07956; Q15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DE 07-MAR-2006, entry version 78.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologues with several different
RT proteins";
RL J. Cell Biol. 103:1635-1648 (1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region";
RL J. Cell Biol. 108:729-736 (1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-Mcutchan F., Framson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins";
RL Biochemistry 25:8418-8425 (1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453 (1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227 (1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
TRP-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;

RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.,
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT module.",
RL J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RT of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hydrazide chemistry, and mass spectrometry.",
RL J. Proteome Res. 4:2070-2080(2005).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-1ib/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04685; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14326; AAA61237.1; ALT_SEQ; mRNA.
DR EMBL; J04835; AAA61178.1; -; Genomic_DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1ZAA; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endoropeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF_3.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR013032; EGF-like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWFC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC; 1; 1.
DR PROSITE; PS01184; VWFC; 2; 1.
KW 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO 0000035842.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...).
FT N-linked (GlcNAc...).
FT C-linked (Man).
FT /FTID=CAR 000205.
FT O-linked (Fuc...).
FT /FTID=CAR 000206.
FT C-linked (Man).
FT /FTID=CAR 000207.
FT C-linked (Man).
FT /FTID=CAR 000208.
FT O-linked (Fuc...).
FT /FTID=CAR 000209.
FT C-linked (Man).
FT /FTID=CAR 000210.
FT O-linked (Fuc...).
FT /FTID=CAR 000211.
FT N-linked (GlcNAc...).
FT N-linked (GlcNAc...).
FT Interchain (Probable).
FT Interchain (Probable).

Query Match 91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.15;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FQGVLANVRFFV 12

Db 208 FQGVLANVRFFV 219

RESULT 5

TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 57.
DE Thrombospondin-1 precursor.
DE Names=Thbs1; Synonyms=Tspl;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RN NUCLEOTIDE SEQUENCE OF 1-490.
RX MEDLINE=90375546; PubMed=2338070;
RA Bornstein P., Alfi D., Devaratyalu S., Framson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
RN [4]
RN PROTEIN SEQUENCE OF 19-37.
RX PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;
RA Chen H., Aeschlimann D., Nowlen J., Mosher D.F.;
RT "Expression and initial characterization of recombinant mouse thrombospondin 1 and thrombospondin 3.";
RL FEBS Lett. 387:36-41(1996).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibrinogen, fibronectin, laminin, type V collagen and integrins alpha-V/beta-1, alpha-V/beta-3 and alpha-IIB/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
EMBL; M62470; AAAS0611.1; -, Genomic DNA.
EMBL; M62450; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62451; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62452; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62453; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62454; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62455; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62456; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62457; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62458; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62459; AAAS0611.1; JOINED; Genomic DNA.
EMBL; M62460; AAAS0611.1; JOINED; Genomic DNA.

DR EMBL; M62461; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62462; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62463; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62464; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62465; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62466; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62467; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62468; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M62469; AAAS0611.1; JOINED; Genomic DNA.
DR EMBL; M87276; AAAS3063.1; -, mRNA.
DR EMBL; J05606; AAAS0611.1; -, Genomic DNA.
DR EMBL; J05605; AAAS0611.1; -, Genomic DNA.
DR PIR; A40558; A40558.
DR HGSP; P07996; ILSL.
DR SMR; P35441; 549-1169.
DR Ensembl; ENSMUSG0000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSPIREPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00266; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01184; VWFC_2; 1.
KW Calcium; Cell adhesion; Direct protein sequencing; EGF-like domain;
KW Glycoprotein; Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTID=PRO_0000035843.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT N-linked (GlcNAc...) (Potential).
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708

FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).

FT DISULFID 270 270 Interchain (Probable).

FT DISULFID 274 274 Interchain (Probable).

FT DISULFID 391 423 By similarity.

FT DISULFID 395 428 By similarity.

FT DISULFID 406 413 By similarity.

FT DISULFID 447 484 By similarity.

FT DISULFID 451 489 By similarity.

FT DISULFID 462 474 By similarity.

FT DISULFID 504 541 By similarity.

FT DISULFID 508 546 By similarity.

FT DISULFID 519 531 By similarity.

FT DISULFID 551 562 By similarity.

FT DISULFID 556 572 By similarity.

FT DISULFID 575 586 By similarity.

FT DISULFID 592 608 By similarity.

FT DISULFID 599 617 By similarity.

FT DISULFID 620 644 By similarity.

FT DISULFID 650 663 By similarity.

FT DISULFID 657 676 By similarity.

FT DISULFID 678 689 By similarity.

FT DISULFID 705 713 By similarity.

FT DISULFID 718 738 By similarity.

FT DISULFID 754 774 By similarity.

FT DISULFID 777 797 By similarity.

FT DISULFID 813 823 By similarity.

FT DISULFID 836 856 By similarity.

FT DISULFID 874 894 By similarity.

FT DISULFID 910 930 By similarity.

FT DISULFID 946 1167 By similarity.

FT CONFLICT 1025 1025 F -> L (in Ref. 2).

SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 91.7%; Score 55; DB 1; Length 1170;

Best Local Similarity 91.7%; Pred. NO. 0.15;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGLANVRVF 12

Db 208 FQGLQNVRFV 219

RESULT 6

Q3TR40 MOUSE

AC Q3TR40 MOUSE PRELIMINARY; PRT; 1170 AA.

DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.

DT 07-FEB-2006, entry version 1.

DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,

DE Clone:A330055N06 product:thrombospondin 1, full insert sequence.

GN Name=Thbs1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

RC MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning."

RL Methods Enzymol. 303:19-44(1999).

RN NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

RC PubMed=16141072; DOI=10.1126/science.1112014;

RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,

RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,

RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,

RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,

RA Ambei-Impicbato A., Apweiler R., Aturaliya R.N., Bailey T.L.,

RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelson J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamaniishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmer S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya S.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki N.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;

RT "The transcriptional landscape of the mammalian genome.";

RL Science 303:1559-1563(2005).

RN NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

RC PubMed=16141073; DOI=10.1126/science.1112009;

RG RIKEN Genome Exploration Research Group, and Genome Science Group

RG (Genome Network Core Team) and the FANTOM Consortium;

RT "Antisense Transcription in the Mammalian Transcriptome.";

RL Science 309:1564-1566(2005).

RN NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;

RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayashizaki Y.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of


```

DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS50026; EGF_3; 2.
DR PROSITE: PS50092; TSP1; 3.
DR PROSITE: PS01208; VWF_C_1; UNKNOWN_1.
DR PROSITE: PS0184; VWF_C_2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.15; Mismatches 0; Gaps 0;
Matches 11; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 FQGVLANRVFV 12
Db 208 FQGVLANRVFV 219
||||| |||||

RESULT 8
Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted under the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

```

CC EMBL: BC050917; AAH50917.1; -; mRNA.
DR HSPF: P07396; 1LSL.
DR SMR: Q80YQ1; 835-1170.
DR Ensembl: ENSMUSG00000040152; Mus musculus.
DR MGI: MGI:98737; Thbs1.
DR GO: GO:0005615; C:extracellular space; RCA.
DR GO: GO:0005615; C:extracellular space; IDA.
DR GO: GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro: IPR013320; Cona_like_subgrp.
DR InterPro: IPR006210; EGF.
DR InterPro: IPR00742; EGF_3.
DR InterPro: IPR001881; EGF_Ca_bd.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR013032; EGF_like_reg.
DR InterPro: IPR003129; Laminin_G_TSP_N.
DR InterPro: IPR008884; TSP1.
DR InterPro: IPR008085; TSP_1.
DR InterPro: IPR003367; tsp_3.
DR InterPro: IPR008859; TSP_C.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS50026; EGF_3; 2.
DR PROSITE: PS50092; TSP1; 3.
DR PROSITE: PS01208; VWF_C_1; UNKNOWN_1.
DR PROSITE: PS0184; VWF_C_2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. No. 0.15; Mismatches 0; Gaps 0;
Matches 11; Conservative 0; Indels 1; Indels 0; Gaps 0;

QY 1 FQGVLANRVFV 12
Db 208 FQGVLANRVFV 219
||||| |||||

RESULT 9
Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q8CGB2
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 Jy9-MC(B) cDNA, RIKEN full-
length enriched library, clone:G930018021 product:thrombospondin 1,
full insert sequence).
DE Name=Thbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor
arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted under the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", *Proc. Natl. Acad. Sci. U.S.A.* 99:16899-16903(2002).

RL [2]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=CZECH II; TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;

RC Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Methods Enzymol. 303:19-44(1999).

RN [4]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;

RX PubMed=1614072; DOI=10.1126/science.1112014;

RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K., Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M., Davis M.J., Wilming L.G., Aidinis V., Allen J.E., Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L., Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M., Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R., Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G., di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G., Fletcher C.P., Fukushima T., Lazarevic D., Lipovich L., Liu J., Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L., Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K., Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P., Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O., Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G., Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M., Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y., Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B., Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K., Tamaoka K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A., Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K., Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C., Grimmond S.M., Tesdale R.D., Liu E.T., Brusic V., Quackenbush J., Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y., Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T., Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N., Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N., Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S., Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J., Hayashizaki Y.;

RT "The transcriptional landscape of the mammalian genome.";

RL Science 309:1559-1563(2005).

RN [5]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;

RX PubMed=1614073; DOI=10.1126/science.1112009;

RG RIKEN Genome Exploration Research Group, and Genome Science Group (Genome Network Core Team) and the FANTOM Consortium;

RT "Antisense Transcription in the Mammalian Transcriptome.";

RL Science 309:1564-1566(2005).

RN [6]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;

RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Tesdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki K., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [7]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schonbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohteki S., Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [8]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [9]

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [10]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RA Aikawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC042422; AAH42422.1; -; mRNA.
DR EMBL; AK145202; BAE26293.1; -; mRNA.
DR HSSP; P07996; ILSL.

Query Match 91.7%; Score 55; DB 2; Length 1171;
Best Local Similarity 91.7%; Pred. No. 0.15;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLVNRFVVF 12
DB 208 FQGVLVNRFVVF 219
|||||
|||||

RESULT 10
TSPI_XENLA STANDARD; PRT; 1173 AA.
AC P15448;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 54.
DE Thrombospondin-1 precursor.
GN Name:thbs1; Synonyms:tspi;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RT "Cloning, characterization and expression of thrombospondin-1 in
RT Xenopus laevis embryos.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-11b/beta-3 (by similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.

```

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
DR HSSP; P07996; ILSL.
DR SMR; P35448; 552-1172.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR01881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR PRINTS; PF00093; VWC; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01844; VWFC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1173 Thrombospondin-1.
FT DOMAIN 23 224 /FTID=PRO_0000035844.
FT DOMAIN 319 376 TSP N-terminal.
FT DOMAIN 382 432 VWFC.
FT DOMAIN 438 493 TSP type-1 1.
FT DOMAIN 495 550 TSP type-1 2.
FT DOMAIN 550 590 TSP type-1 3.
FT DOMAIN 591 648 EGF-like 1.
FT DOMAIN 649 693 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 726 761 EGF-like 3.
FT DOMAIN 762 784 TSP type-3 1.
FT DOMAIN 785 820 TSP type-3 2.
FT DOMAIN 821 843 TSP type-3 3.
FT DOMAIN 844 881 TSP type-3 4.
FT DOMAIN 882 917 TSP type-3 5.
FT DOMAIN 918 953 TSP type-3 6.
FT DOMAIN 954 1173 TSP type-3 7.
FT REGION 23 235 Heparin-binding (Potential).
FT MOTIF 929 931 Cell attachment site (Potential).
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 363 363 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 705 705 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 711 711 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1070 1070 N-linked (GlcNAc...) (Potential).
FT DISULFID 394 426 By similarity.
FT DISULFID 398 431 By similarity.
FT DISULFID 409 416 By similarity.
FT DISULFID 450 487 By similarity.
FT DISULFID 454 492 By similarity.
FT DISULFID 465 477 By similarity.
FT DISULFID 507 544 By similarity.
FT DISULFID 511 549 By similarity.
FT DISULFID 522 534 By similarity.

```

```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 602 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 91.7%; Score 55; DB 1; Length 1173;
Best Local Similarity 91.7%; Pred. No. 0.15;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 12
DB 211 FQGVLANVRVF 222

RESULT 11
Q59E99 HUMAN
ID Q59E99_HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AB209912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939, 889-1225.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0003198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
```

```
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF000008; EGF_2.
DR Pfam; PF000090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16E57157B12 CRC64;

Query Match 91.7%; Score 55; DB 2; Length 1225;
Best Local Similarity 91.7%; Pred. No. 0.16;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 12
DB 263 FQGVLANVRVF 274

RESULT 12
Q5SPG5 BRARE
ID Q5SPG5_BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n (Fragment).
GN ORFNames=DKFZ-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AL928866; CAI20599.1; -; Genomic_DNA.
DR SMR; Q5SPG5; 751-804, 754-1089.
DR Ensembl; ENSDARG0000010785; Danio rerio.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0003198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008859; TSP_1.
```

```
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP 1; 2.
DR Pfam; PF02412; TSP 3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00018; EF_HAND 1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS01208; VWFC 1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
KW Cell adhesion; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;

Query Match 83.3%; Score 50; DB 2; Length 1090;
Best Local Similarity 83.3%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
Db 185 FMGVLANVRVVF 196

RESULT 13
ID Q4S758.TETNG PRELIMINARY; PRT; 1193 AA.
AC Q4S758;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, entry version 8.
DE Chromosome 14 SCAF14723, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00022976001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segrens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
```

```
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAAE01014723; CAG03524.1; -; Genomic_DNA.
DR SMR; Q4S758; 811-1148.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005195; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF 3.
DR InterPro; IPR01881; EGF_Ca bd.
DR InterPro; IPR013032; EGF_like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP 1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF 2; UNKNOWN_1.
DR PROSITE; PS00026; EGF 3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC 1; 1.
DR PROSITE; PS0184; VWFC_2; 1.
KW Cell adhesion.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;

Query Match 83.3%; Score 50; DB 2; Length 1193;
Best Local Similarity 83.3%; Pred. No. 1.6;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
Db 187 FMGVLANVRVVF 198

RESULT 14
ID Q5U903.PIG PRELIMINARY; PRT; 249 AA.
AC Q5U903;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Thrombospondin 1 (Fragment).
GN Name=Thbs1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
NUCLEOTIDE SEQUENCE.
RA Zhang K., Maucio G., Hauet T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AV773342; AAV38110.1; -; mRNA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR001007; VWF_C.
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 23.8966 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-29

Perfect score: 60

Sequence: 1 FQGVLANVRVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5 COMB pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6 COMB pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7 COMB pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H COMB pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS COMB pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE COMB pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	91.7	825	2	US-09-949-002-482
2	55	91.7	831	2	US-09-939-853A-97
3	55	91.7	831	2	US-09-939-853A-98
4	55	91.7	1170	1	US-08-313-288B-20
5	55	91.7	1170	2	US-09-657-472-2
6	55	91.7	1170	2	US-09-949-002-350
7	38	63.3	175	2	US-09-230-637-24
8	38	63.3	175	2	US-10-003-632C-5
9	37	61.7	326	2	US-09-083-268-17
10	37	61.7	326	2	US-08-981-998A-19
11	37	61.7	418	2	US-09-648-281-12
12	37	61.7	418	2	US-09-083-268-5
13	37	61.7	740	2	US-09-134-000C-6441
14	37	61.7	1045	2	US-09-949-016-1112
15	37	61.7	1135	2	US-08-981-998A-5
16	37	61.7	1172	1	US-08-313-288B-19
17	37	61.7	1172	2	US-09-949-016-6333
18	36	60.0	61	2	US-09-270-767-61396
19	36	60.0	304	2	US-09-270-767-45864
20	35	58.3	365	2	US-09-902-540-10979
21	35	58.3	413	2	US-09-252-991A-20866
22	35	58.3	1288	2	US-08-762-428A-6
23	34	56.7	273	2	US-09-252-991A-20770
24	34	56.7	343	1	US-08-454-196-7
25	34	56.7	343	1	US-08-286-819A-8
26	34	56.7	343	2	US-08-980-357-8

27	34	56.7	343	2	US-09-064-033-7	Sequence 7, Appli
28	34	56.7	343	2	US-09-291-046-7	Sequence 7, Appli
29	34	56.7	343	2	US-09-357-373-8	Sequence 8, Appli
30	34	56.7	346	2	US-09-083-268-16	Sequence 16, Appl
31	34	56.7	346	2	US-08-981-998A-16	Sequence 16, Appl
32	34	56.7	390	2	US-08-650-766-7	Sequence 7, Appli
33	34	56.7	390	2	US-08-922-635-6	Sequence 6, Appli
34	34	56.7	390	2	US-09-389-487-7	Sequence 7, Appli
35	34	56.7	390	2	US-09-414-643-6	Sequence 6, Appli
36	34	56.7	455	2	US-09-543-681A-8288	Sequence 8288, Ap
37	34	56.7	559	2	US-09-364-206-47	Sequence 47, Appl
38	34	56.7	651	2	US-08-650-766-6	Sequence 6, Appli
39	34	56.7	651	2	US-08-922-635-5	Sequence 5, Appli
40	34	56.7	651	2	US-09-389-487-6	Sequence 6, Appli
41	34	56.7	651	2	US-09-414-643-5	Sequence 5, Appli
42	34	56.7	710	2	US-09-489-039A-14121	Sequence 14121, A
43	34	56.7	1070	2	US-08-922-635-22	Sequence 22, Appl
44	34	56.7	1070	2	US-09-414-643-22	Sequence 22, Appl
45	34	56.7	1312	2	US-09-041-886-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION

; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,002

; CURRENT FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/231,401

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 482

; LENGTH: 825

; TYPE: PRT

; ORGANISM: Human

US-09-949-002-482

Query Match 91.7%; Score 55; DB 2; Length 825;
Best Local Similarity 91.7%; Pred. No. 0.026;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVF 12

||||| |||||

Db 308 FQGVLANVRVF 319

RESULT 2

US-09-939-853A-97

; Sequence 97, Application US/09939853A

; Patent No. 6989232

; GENERAL INFORMATION:

; APPLICANT: Burgess et al.

; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-099

; CURRENT APPLICATION NUMBER: US/09/939,853A

; CURRENT FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: 60/228,191

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: 60/267,300

; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/269,961

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/277,337

; PRIOR FILING DATE: 2001-03-20

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.7%; Score 55; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLANVRVF 12
      ||||| |||||
Db      208 FQGVLANVRVF 219

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/093939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match          91.7%; Score 55; DB 2; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLANVRVF 12
      ||||| |||||
Db      208 FQGVLANVRVF 219

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLANVRVF 12
      ||||| |||||
Db      208 FQGVLANVRVF 219

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          91.7%; Score 55; DB 2; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 FQGVLANVRVF 12
      ||||| |||||
Db      208 FQGVLANVRVF 219

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```

; SOFTWARE: PatentIn Ver 3.1
; SEQ ID NO 5
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-632C-5

Query Match      63.3%; Score 38; DB 2; Length 175;
Best Local Similarity 70.0%; Pred. No. 10;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 FQGVLANRVF 10
      | : : |||||
Db      51 FEAWLANRVF 60

RESULT 9
US-09-083-268-17
; Sequence 17, Application US/09083268
; Patent No. 6673535
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; TITLE OF INVENTION: ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.
; STREET: 119 No. 6673535th Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McCormack, Myra H
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 232.00010101
; TELEPHONE: 612/305-1220
; TELEFAX: 612/305-1228
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
US-09-083-268-17

Query Match      61.7%; Score 37; DB 2; Length 326;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 FQGVLANRVF 11
      | : : |||||
Db      85 FDGIYANRVF 95

RESULT 10
US-08-981-998A-17
; Sequence 17, Application US/08981998A
; Patent No. 6844431

```

GENERAL INFORMATION:
APPLICANT: PULST, STEFAN M.
TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR ATAXIA-2 AND PRODUCTS RELATED THERETO
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: MUEITING, RAASCH & GEBHARDT, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: 55401
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/981,998A
FILING DATE: 11-May-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 97/42314
FILING DATE: 08-MAY-1997
APPLICATION NUMBER: US 08/727,084
FILING DATE: 08-OCT-1996
APPLICATION NUMBER: US 60/022,207
FILING DATE: 19-JUL-1996
APPLICATION NUMBER: US 60/017,368
FILING DATE: 08-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 232.00010120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-08-981-998A-17

Query Match 61.7%; Score 37; DB 2; Length 326;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVLNVRV 11
DB 85 FDGIYANVRV 95

RESULT 11
US-09-648-281-12
Query Match 61.7%; Score 37; DB 2; Length 326;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Pulst, Stefan M.
TITLE OF INVENTION: Transgenic Animal Model of Neurodegenerative Disease and Methods of Use
FILE REFERENCE: P-CE 4336
CURRENT APPLICATION NUMBER: US/09/648,281
CURRENT FILING DATE: 2000-08-24
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 418
TYPE: PRT
ORGANISM: mus musculus

US-09-648-281-12

Query Match 61.7%; Score 37; DB 2; Length 418;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVLNVRV 11
DB 85 FDGIYANVRV 95

RESULT 12
US-09-083-268-5
Query Match 61.7%; Score 37; DB 2; Length 418;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Pulst, Stefan M.
TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR ATAXIA-2 AND PRODUCTS RELATED THERETO
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueiting, Raasch & Gebhardt, P.A.
STREET: 119 No. 6673535th Fourth Street
CITY: Minneapolis
STATE: Minnesota
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/727,084
FILING DATE: 08-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 232.00010101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/305-1220
TELEFAX: 612/305-1228
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-083-268-5

Query Match 61.7%; Score 37; DB 2; Length 418;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 FOGVLNVRV 11
DB 85 FDGIYANVRV 95

RESULT 13
US-09-134-000C-6441
Query Match 61.7%; Score 37; DB 2; Length 418;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C

```
;
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6441
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (11)-(11)
; OTHER INFORMATION: Amino acid 11 is Xaa wherein Xaa = any amino acid.
US-09-134-000C-6441

Query Match          61.7%; Score 37; DB 2; Length 740;
Best Local Similarity 41.7%; Pred. No. 74;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 FQGVLANVRVVF 12
Db      728 FCGILGNINFIY 739

RESULT 14
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11112

Query Match          61.7%; Score 37; DB 2; Length 1045;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 FQGVLANVRVVF 12
Db      281 FRGLQNHLVF 292

RESULT 15
US-08-981-998A-5
; Sequence 5, Application US/08981998A
; Patent No. 6844431
; GENERAL INFORMATION:
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERECELLAR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
```

```
;
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/981,998A
; APPLICATION NUMBER: US/08/981,998A
; FILING DATE: 11-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/42314
; FILING DATE: 08-MAY-1997
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: US 60/022,207
; FILING DATE: 19-JUL-1996
; APPLICATION NUMBER: US 60/017,388
; FILING DATE: 08-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MUETING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-981-998A-5

Query Match          61.7%; Score 37; DB 2; Length 1135;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 FQGVLANVRVVF 11
Db      79 FDGIYANVRMV 89

Search completed: June 5, 2006, 22:48:55
Job time : 23.8966 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 78.6207 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-29
Perfect score: 60
Sequence: 1 FQGVLANVRFPV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	91.7	12	4 US-10-474-213-28	Sequence 28, Appl
2	55	91.7	240	4 US-10-419-462-40	Sequence 40, Appl
3	55	91.7	240	5 US-10-782-968-40	Sequence 40, Appl
4	55	91.7	432	5 US-10-741-600-1020	Sequence 1020, Ap
5	55	91.7	432	5 US-10-741-600-1022	Sequence 1022, Ap
6	55	91.7	459	6 US-11-043-806-462	Sequence 462, App
7	55	91.7	466	3 US-09-925-301-1047	Sequence 1047, App
8	55	91.7	555	6 US-11-043-806-454	Sequence 454, App
9	55	91.7	578	6 US-11-043-806-456	Sequence 456, App
10	55	91.7	685	6 US-11-043-806-452	Sequence 452, App
11	55	91.7	804	6 US-11-043-806-453	Sequence 453, App
12	55	91.7	828	6 US-11-043-806-455	Sequence 455, App
13	55	91.7	831	3 US-09-939-853A-97	Sequence 97, Appl
14	55	91.7	831	3 US-09-939-853A-98	Sequence 98, Appl
15	55	91.7	855	6 US-11-043-806-461	Sequence 461, App
16	55	91.7	1000	6 US-11-043-806-457	Sequence 457, App
17	55	91.7	1105	6 US-11-043-806-458	Sequence 458, App
18	55	91.7	1150	4 US-10-296-733-1	Sequence 1, Appli
19	55	91.7	1152	3 US-09-919-603-1	Sequence 1, Appli
20	55	91.7	1169	5 US-10-317-821B-7	Sequence 7, Appli
21	55	91.7	1170	4 US-10-020-141-12	Sequence 12, Appl
22	55	91.7	1170	4 US-10-017-721-2	Sequence 2, Appli
23	55	91.7	1170	4 US-10-021-660-114	Sequence 114, App
24	55	91.7	1170	4 US-10-008-093-2	Sequence 2, Appli
25	55	91.7	1170	4 US-10-295-027-1170	Sequence 1170, Ap
26	55	91.7	1170	4 US-10-211-462-38	Sequence 38, Appl
27	55	91.7	1170	4 US-10-231-956A-482	Sequence 482, App

28	55	91.7	1170	4 US-10-419-462-38	Sequence 38, Appl
29	55	91.7	1170	5 US-10-741-600-1018	Sequence 1018, Ap
30	55	91.7	1170	5 US-10-741-600-1019	Sequence 1019, Ap
31	55	91.7	1170	5 US-10-741-600-1021	Sequence 1021, Ap
32	55	91.7	1170	5 US-10-782-968-38	Sequence 38, Appl
33	55	91.7	1170	5 US-10-849-989-44	Sequence 44, Appl
34	55	91.7	1170	5 US-10-631-467-548	Sequence 548, App
35	55	91.7	1170	5 US-10-631-467-1376	Sequence 1376, Ap
36	55	91.7	1170	5 US-10-831-997-2	Sequence 2, Appli
37	55	91.7	1170	5 US-10-995-561-594	Sequence 594, App
38	55	91.7	1170	5 US-10-995-561-595	Sequence 595, App
39	55	91.7	1170	5 US-10-995-561-596	Sequence 596, App
40	55	91.7	1170	6 US-11-037-713-51	Sequence 51, Appl
41	55	91.7	1170	6 US-11-046-644-28	Sequence 28, Appl
42	55	91.7	1170	6 US-11-046-456-28	Sequence 28, Appl
43	39	65.0	15	4 US-10-285-394-153	Sequence 153, App
44	38	63.3	86	4 US-10-437-963-176780	Sequence 176780,
45	38	63.3	175	4 US-10-003-632C-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-10-474-213-28
; Sequence 28, Application US/10474213
; Publication No. US20040214248A1
; GENERAL INFORMATION:
; APPLICANT: Roberts, David D
; APPLICANT: Kruszsch, Henry C
; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF
; FILE REFERENCE: 224329
; CURRENT APPLICATION NUMBER: US/10/474,213
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/10535
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 60/281,994
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-474-213-28

Query Match 91.7%; Score 55; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00076;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRFPV 12
| | | | | | | | | |
Db 1 FQGVLANVRFPV 12

RESULT 2
US-10-419-462-40
; Sequence 40, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40

; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match 91.7%; Score 55; DB 4; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.022; 1; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 190 FQGVLANVRVFV 201
||||| |||||

RESULT 3

US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match 91.7%; Score 55; DB 5; Length 240;
Best Local Similarity 91.7%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 190 FQGVLANVRVFV 201
||||| |||||

RESULT 4

US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 91.7%; Score 55; DB 5; Length 432;

Best Local Similarity 91.7%; Pred. No. 0.042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219
||||| |||||

RESULT 5

US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 91.7%; Score 55; DB 5; Length 432;
Best Local Similarity 91.7%; Pred. No. 0.042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219
||||| |||||

RESULT 6

US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Method
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match 91.7%; Score 55; DB 6; Length 459;
Best Local Similarity 91.7%; Pred. No. 0.045;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219
||||| |||||

RESULT 7

US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

;
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL06
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 91.7%; Score 55; DB 3; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.046;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
| | | | | | | | | |
DB 261 FQGVLANVRVVF 272

RESULT 8

US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 91.7%; Score 55; DB 6; Length 555;
Best Local Similarity 91.7%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
| | | | | | | | | |
DB 208 FQGVLANVRVVF 219

RESULT 9

US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 91.7%; Score 55; DB 6; Length 578;
Best Local Similarity 91.7%; Pred. No. 0.058;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
| | | | | | | | | |
DB 208 FQGVLANVRVVF 219

RESULT 10

US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 91.7%; Score 55; DB 6; Length 685;
Best Local Similarity 91.7%; Pred. No. 0.07;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
| | | | | | | | | |
DB 208 FQGVLANVRVVF 219

RESULT 11

US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 91.7%; Score 55; DB 6; Length 804;
Best Local Similarity 91.7%; Pred. No. 0.084;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVVF 12
| | | | | | | | | |
DB 208 FQGVLANVRVVF 219

RESULT 12

US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match          91.7%; Score 55; DB 6; Length 828;
Best Local Similarity 91.7%; Pred. No. 0.087;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match          91.7%; Score 55; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.087;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

US-09-939-853A-98

```
Query Match          91.7%; Score 55; DB 3; Length 831;
Best Local Similarity 91.7%; Pred. No. 0.087;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219
```

RESULT 15

```
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; TITLE OF INVENTION: thereof for Diagnosis of Prostate Cancer
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461
```

```
Query Match          91.7%; Score 55; DB 6; Length 855;
Best Local Similarity 91.7%; Pred. No. 0.09;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 FQGVLANVRVFV 12
Db 208 FQGVLANVRVFV 219
```

Search completed: June 6, 2006, 00:00:11
Job time : 79.6207 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.72414 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-29

Perfect score: 60

Sequence: 1 QGVLANVRVVF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	56.7	331	6	US-10-953-349-34674
2	34	56.7	358	6	US-10-953-349-34673
3	34	56.7	382	6	US-10-953-349-34672
4	34	56.7	1504	6	US-10-505-928-662
5	33	55.0	182	6	US-10-953-349-15318
6	33	55.0	186	6	US-10-953-349-27518
7	33	55.0	188	6	US-10-953-349-27517
8	33	55.0	213	6	US-10-953-349-27516
9	33	55.0	297	7	US-11-293-697-4132
10	32	53.3	166	6	US-10-953-349-36110
11	32	53.3	173	6	US-10-953-349-36109
12	32	53.3	209	6	US-10-953-349-36108
13	32	53.3	249	6	US-10-953-349-11050
14	32	53.3	254	6	US-10-953-349-11049
15	32	53.3	286	6	US-10-953-349-11048
16	31	51.7	210	6	US-10-953-349-21828
17	31	51.7	223	6	US-10-953-349-21827
18	31	51.7	272	6	US-10-953-349-21826
19	31	51.7	392	6	US-10-196-749-160
20	31	51.7	407	6	US-10-953-349-26200
21	31	51.7	461	6	US-10-505-928-282
22	31	51.7	980	7	US-11-242-505A-36
23	30	50.0	164	6	US-10-953-349-1159
24	30	50.0	249	6	US-10-953-349-1158
25	30	50.0	250	6	US-10-953-349-1157

26	30	50.0	315	6	US-10-953-349-9204	Sequence 9204, Ap
27	30	50.0	362	6	US-10-953-349-35080	Sequence 35080, A
28	30	50.0	400	6	US-10-953-349-9203	Sequence 9203, Ap
29	30	50.0	401	6	US-10-953-349-9202	Sequence 9202, Ap
30	30	50.0	418	6	US-10-953-349-35079	Sequence 35079, A
31	30	50.0	437	6	US-10-504-120-32	Sequence 32, Appl
32	30	50.0	445	6	US-10-953-349-2333	Sequence 2333, Ap
33	30	50.0	467	6	US-10-953-349-2332	Sequence 2332, Ap
34	30	50.0	516	6	US-10-953-349-2331	Sequence 2331, Ap
35	30	50.0	843	6	US-10-953-349-16527	Sequence 16527, A
36	30	50.0	962	6	US-10-953-349-16526	Sequence 16526, A
37	30	50.0	1014	6	US-10-953-349-16525	Sequence 16525, A
38	29.5	49.2	157	6	US-10-953-349-6894	Sequence 6894, Ap
39	29.5	49.2	208	6	US-10-953-349-6893	Sequence 6893, Ap
40	29.5	49.2	263	6	US-10-953-349-6892	Sequence 6892, Ap
41	29	48.3	114	6	US-10-953-349-9965	Sequence 9965, Ap
42	29	48.3	130	6	US-10-953-349-29183	Sequence 29183, A
43	29	48.3	143	6	US-10-953-349-9964	Sequence 9964, Ap
44	29	48.3	143	6	US-10-953-349-29182	Sequence 29182, A
45	29	48.3	156	6	US-10-953-349-25905	Sequence 25905, A

ALIGNMENTS

RESULT 1

US-10-953-349-34674
; Sequence 34674, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match 56.7%; Score 34; DB 6; Length 331;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	2	QGVLANVRV 11
DB	225	QGVLFNIQIV 234

RESULT 2

US-10-953-349-34673
; Sequence 34673, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

Query Match 56.7%; Score 34; DB 6; Length 358;

```
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15318
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15318

Query Match      55.0%; Score 33; DB 6; Length 182;
Best Local Similarity 60.0%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 GVLNVRFV 12
       ||| |||
Db     104 GLLMNNFV 113
       ||| |||

RESULT 6
US-10-953-349-27518
; Sequence 27518, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27518
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-27518

Query Match      55.0%; Score 33; DB 6; Length 186;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 FQGVLANRVFV 11
       | ||| |||
Db     63 FSAALAEVRFV 73
       ||| |||

RESULT 7
US-10-953-349-27517
; Sequence 27517, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27517
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-27517

Query Match      55.0%; Score 33; DB 6; Length 188;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 FQGVLANRVFV 11
       | ||| |||

; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15318
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15318

Query Match      56.7%; Score 34; DB 6; Length 382;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 QGVLANRVFV 11
       ||| ||| |||
Db     276 QGVLFNIQYV 285
       ||| ||| |||

RESULT 4
US-10-505-928-662
; Sequence 662, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 662
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-662

Query Match      56.7%; Score 34; DB 6; Length 1504;
Best Local Similarity 54.5%; Pred. No. 63;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 QGVLANRVFV 12
       | ||| ||| |||
Db     751 QHILSLRFV 761
       ||| ||| ||| |||

RESULT 5
US-10-953-349-15318
; Sequence 15318, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
```

```
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15318
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15318

Query Match      55.0%; Score 33; DB 6; Length 182;
Best Local Similarity 60.0%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 GVLNVRFV 12
       ||| |||
Db     104 GLLMNNFV 113
       ||| |||

RESULT 6
US-10-953-349-27518
; Sequence 27518, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34672
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34672

Query Match      56.7%; Score 34; DB 6; Length 382;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 QGVLANRVFV 11
       ||| ||| |||
Db     276 QGVLFNIQYV 285
       ||| ||| |||

RESULT 4
US-10-505-928-662
; Sequence 662, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 662
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-662

Query Match      56.7%; Score 34; DB 6; Length 1504;
Best Local Similarity 54.5%; Pred. No. 63;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 QGVLANRVFV 12
       | ||| ||| |||
Db     751 QHILSLRFV 761
       ||| ||| ||| |||

RESULT 5
US-10-953-349-15318
; Sequence 15318, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
```

Db 65 FSAALAEVRV 75

RESULT 8

US-10-953-349-27516
; Sequence 27516, Application US/10953349
; Publication No. US20060107345A1

GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3

SEQ ID NO 27516

LENGTH: 213

TYPE: PRT

ORGANISM: Triticum aestivum

US-10-953-349-27516

Query Match 55.0%; Score 33; DB 6; Length 213;

Best Local Similarity 63.6%; Pred. No. 12;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 FQGVLANVRV 11

Db 90 FSAALAEVRV 100

RESULT 9

US-11-293-697-4132
; Sequence 4132, Application US/11293697
; Publication No. US20060105376A1

GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4132

LENGTH: 297

TYPE: PRT

ORGANISM: Homo sapiens

US-11-293-697-4132

Query Match 55.0%; Score 33; DB 7; Length 297;

Best Local Similarity 70.0%; Pred. No. 17;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVLNVRV 12

Db 234 GVFNVRV 243

RESULT 10

US-10-953-349-36110
; Sequence 36110, Application US/10953349
; Publication No. US20060107345A1

GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3

SEQ ID NO 36110

LENGTH: 166

TYPE: PRT

ORGANISM: Zea mays subsp. mays

US-10-953-349-36110

Query Match 53.3%; Score 32; DB 6; Length 173;

Best Local Similarity 85.7%; Pred. No. 15;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLAN 7

Db 166 FQGVAN 172

RESULT 12

US-10-953-349-36108

Query Match

53.3%; Score 32; DB 6; Length 173;

Best Local Similarity 85.7%; Pred. No. 15;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLAN 7

Db 166 FQGVAN 172

; Sequence 36108, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36108
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30)..(30)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (38)..(38)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (52)..(52)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (86)..(86)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (88)..(88)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-36108

Query Match 53.3%; Score 32; DB 6; Length 209;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLAN 7
:|||||:
Db 202 FQGVVAN 208

RESULT 13
US-10-953-349-11050
; Sequence 11050, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11050
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11050

Query Match 53.3%; Score 32; DB 6; Length 249;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 VLANRVFV 11
:|||||:
Db 238 ILANVKFI 245

RESULT 14
US-10-953-349-11049
; Sequence 11049, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11049
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11049

Query Match 53.3%; Score 32; DB 6; Length 254;
Best Local Similarity 62.5%; Pred. No. 23;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 VLANRVFV 11
:|||||:
Db 243 ILANVKFI 250

RESULT 15
US-10-953-349-11048
; Sequence 11048, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11048
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11048

Query Match 53.3%; Score 32; DB 6; Length 286;
Best Local Similarity 62.5%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 VLANRVFV 11
:|||||:
Db 275 ILANVKFI 282

Search completed: June 6, 2006, 00:12:56
Job time : 3.82414 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:08:53 ; Search time 83.5431 Seconds
(without alignments)
60.201 Million cell updates/sec

Title: US-10-030-735-30
Perfect score: 55
Sequence: 1 FQVQLQVRFV 11

Scoring table: BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*
10: Geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	11	4 AAB35360	Aab35360 Alpha3bet
2	55	100.0	12	4 AAB35352	Aab35352 Alpha3bet
3	55	100.0	12	4 AAB35366	Aab35366 Alpha3bet
4	55	100.0	12	4 AAB35378	Aab35378 Alpha3bet
5	55	100.0	12	6 ABG72834	Abg72834 Thrombos
6	55	100.0	240	8 ADL70641	Adl70641 Human thr
7	55	100.0	432	8 ADQ39359	Adq39359 Human myo
8	55	100.0	432	8 ADQ39357	Adq39357 Human myo
9	55	100.0	459	4 AAU02916	Aau02916 Angiotens
10	55	100.0	466	3 AAB43602	Aab43602 Human can
11	55	100.0	546	4 AAU02915	Aau02915 Angiotens
12	55	100.0	548	7 ADN02474	Adn02474 TSF poly
13	55	100.0	555	4 AAU02914	Aau02914 Angiotens
14	55	100.0	731	4 AAU02913	Aau02913 Angiotens
15	55	100.0	1152	3 AAB00042	Aab00042 Human thr
16	55	100.0	1152	5 AAU74771	Aau74771 Human thr
17	55	100.0	1152	5 AAB82285	Aab82285 Human thr
18	55	100.0	1170	4 AAB74450	Aab74450 Human var
19	55	100.0	1170	4 AAB90800	Aab90800 Human she
20	55	100.0	1170	5 AAE25030	Aae25030 Human thr
21	55	100.0	1170	5 AAU75315	Aau75315 Human thr
22	55	100.0	1170	5 ABP96780	Abp96780 Human COP
23	55	100.0	1170	6 ABU03474	Abu03474 Angiogene

24	55	100.0	1170	6 ABG74673	Abg74673 Human THB
25	55	100.0	1170	6 AAE36228	Aae36228 Human THB
26	55	100.0	1170	7 ABR62059	Abr62059 Human thr
27	55	100.0	1170	7 ADN39852	Adn39852 Cancer/ge
28	55	100.0	1170	8 ADJ76124	Adj76124 Marker ge
29	55	100.0	1170	8 ADJ75296	Adj75296 Marker ge
30	55	100.0	1170	8 ADL70639	Adl70639 Human thr
31	55	100.0	1170	8 ADL35874	Adl35874 Human thr
32	55	100.0	1170	8 ADQ26070	Adq26070 Thrombos
33	55	100.0	1170	8 ADF54179	Adf54179 Human PRO
34	55	100.0	1170	8 ADQ39358	Adq39358 Human myo
35	55	100.0	1170	8 ADQ39356	Adq39356 Human myo
36	55	100.0	1170	8 ADQ39355	Adq39355 Human myo
37	55	100.0	1170	9 AD221688	Ad221688 Thrombos
38	55	100.0	1170	9 AEB87781	Aeb87781 Human thr
39	55	100.0	1170	9 AEB46751	Aeb46751 Human thr
40	52	94.5	12	4 AAB35373	Aab35373 Alpha3bet
41	52	94.5	12	4 AAB35381	Aab35381 Alpha3bet
42	51	92.7	10	4 AAB35355	Aab35355 Alpha3bet
43	51	92.7	12	4 AAB35364	Aab35364 Alpha3bet
44	51	92.7	12	4 AAB35374	Aab35374 Alpha3bet
45	51	92.7	12	4 AAB35347	Aab35347 Alpha3bet

ALIGNMENTS

RESULT 1
AAB35360
ID AAB35360 standard; peptide; 11 AA.
XX
AC AAB35360;
XX
DT 08-MAY-2001 (first entry)
XX
DE Alpha3bet1 integrin binding peptide #25.
XX
KW Alpha3bet1 integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
PN WO200105812-A2.
XX
PD 25-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US018986.
XX
PR 15-JUL-1999; 99US-0144549P.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
Roberts DD, Kruttsch HC;
XX
WPI; 2001-182656/18.
XX
New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
Claim 4; Page 34; 84pp; English.
XX
The present invention provides a number of peptides which bind to
CC alpha3bet1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
XX of the invention
XX
SQ Sequence 11 AA;

```

Query Match      100.0%; Score 55; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00076; Mismatches 0; Gaps 0;
Matches 11; Conservative 0; Indels 0;

QY 1 FQGVQLQNVRFV 11
   |||||
DB 1 FQGVQLQNVRFV 11

RESULT 2
AAB35352
ID AAB35352 standard; peptide; 12 AA.
AC AAB35352;
XX
XX 08-MAY-2001 (first entry)
XX
XX Alpha3betal integrin binding peptide #17.
XX
XX Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
XX WO200105812-A2.
XX
XX 25-JAN-2001.
XX
XX 12-JUL-2000; 2000WO-US018986.
XX
XX 15-JUL-1999; 99US-0144549P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Roberts DD, Krutzsch HC;
XX
XX WPI; 2001-182656/18.
XX
XX New peptides that bind to or are recognized by alpha3-betal integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
XX Claim 4; Page 34; 84pp; English.
XX
XX The present invention provides a number of peptides which bind to
CC alpha3betal integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
XX Query Match      100.0%; Score 55; DB 4; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 0.00084;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
DB 1 FQGVQLQNVRFV 11

RESULT 4
AAB35378
ID AAB35378 standard; peptide; 12 AA.
XX
XX AAB35378;
XX
XX 08-MAY-2001 (first entry)
XX
XX Alpha3betal integrin binding peptide #43.
XX
XX Alpha3betal integrin; angiogenesis; cell proliferation; cancer;
KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;
KW macular degeneration; psoriasis; cell adhesion; cell motility.
XX
OS Synthetic.
XX
XX WO200105812-A2.
XX
XX 25-JAN-2001.
XX
XX 12-JUL-2000; 2000WO-US018986.
XX
XX 15-JUL-1999; 99US-0144549P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Roberts DD, Krutzsch HC;

QY 1 FQGVQLQNVRFV 11
   |||||
DB 1 FQGVQLQNVRFV 11

RESULT 3
AAB35366
ID AAB35366 standard; peptide; 12 AA.
XX
XX AAB35366;
XX
XX 08-MAY-2001 (first entry)
XX
XX Alpha3betal integrin binding peptide #31.

```


XX WPI; 2001-182656/18.
XX
XX New peptides that bind to or are recognized by alpha3-beta1 integrins,
PT useful for inhibiting cell adhesion to extracellular matrix, cell
PT motility and proliferation and for treating rheumatoid arthritis and
PT cancer.
XX
XX Example 2; Page 34; 84pp; English.
XX
XX The present invention provides a number of peptides which bind to
CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion
CC and motility, and in the treatment of cancer, diabetic retinopathy,
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis
CC and restenosis. The present sequence is an example of one of the peptides
CC of the invention
XX
XX Sequence 12 AA;
SQ

Query Match 100.0%; Score 55; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQLQNVRFV 11
Db 1 FQGVQLQNVRFV 11

RESULT 5
ABG72834
ID ABG72834 standard; peptide; 12 AA.
XX
XX ABG72834;
AC
XX
XX 24-FEB-2003 (first entry)
DT
XX
XX Thrombospondin-1 sequence containing synthetic peptide.
DE
XX
XX Human; thrombospondin-1; cytostatic; immunostimulant; cancer;
KW epithelial cancer; lung cancer; papillary renal cell carcinoma;
KW colon cancer; small-cell lung cancer; SCLC; melanoma.
XX
XX Synthetic.
OS
XX
XX WO200281630-A2.
PN
XX
XX 17-OCT-2002.
PD
XX
XX 03-APR-2002; 2002WO-US010535.
PF
XX
XX 06-APR-2001; 2001US-0281994P.
PR
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Roberts DD, Krutzsch HC;
PI
XX
XX WPI; 2003-103329/09.
DR
XX
XX A new diagnosis for cancer other than prostate cancer in a mammal useful
PT to detect cancer including lung cancer, particularly small cell lung
PT cancer and melanoma comprises detecting semenogelin in a sample.
XX
XX Example 1; Page 14; 32pp; English.
PS
XX
XX The invention relates to diagnosing cancer other than prostate cancer in
CC a male mammal, comprising assaying a test sample for increased level of
CC semenogelin, or cancer in a female by assaying for the presence of
CC semenogelin. Administering a semenogelin protein or polypeptide fragment
CC or a semenogelin-specific antibody or active fragment, or a recombinant
CC vector expressing the protein or antibody, is useful for inducing an
CC immune response to a cancer in a mammal, where the cancer is not prostate
CC cancer and semenogelin is a marker. The invention is used to diagnose
CC cancer, particularly of epithelial origin such as lung cancer, papillary

CC renal cell carcinoma, colon cancer, especially small-cell lung cancer
CC (SCLC), or a melanoma. The present sequence represents the amino acid
CC sequence of the thrombospondin-1 sequence containing synthetic peptide
CC which binds to alpha-3-beta-1 integrin
XX
XX Sequence 12 AA;
SQ

Query Match 100.0%; Score 55; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQLQNVRFV 11
Db 1 FQGVQLQNVRFV 11

RESULT 6
ADL70641
ID ADL70641 standard; protein; 240 AA.
XX
XX ADL70641;
AC
XX
XX 20-MAY-2004 (first entry)
DT
XX
XX Human thrombospondin-1 N-terminal domain.
DE
XX
XX Human; thrombospondin-1; epitope; cancer; diagnosis.
KW
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 23..32
FT /note= "Heparin binding region"
FT Region 77..82
FT /note= "Heparin binding region"
FT Region 151..164
FT /note= "Fibrinogen binding region"
XX
XX WO2004018995-A2.
PN
XX
XX 04-MAR-2004.
PD
XX
XX 20-AUG-2003; 2003WO-US026023.
PF
XX
XX 23-AUG-2002; 2002US-0405494P.
PR
XX
XX 21-APR-2003; 2003US-00419462.
PR
XX
XX (WILL/) WILLIAMS K J.
PA
XX
XX Williams KJ;
PI
XX
XX WPI; 2004-226901/21.
DR
XX
XX New purified thrombospondin fragment extracted from a body fluid, useful
PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
PT or leukemia or as calibrators, indicators, immunogens and analytes.
XX
XX Disclosure; SEQ ID NO 40; 76pp; English.
PS
XX
XX The present sequence is that of the N-terminal domain of human
CC thrombospondin-1 (TSP) ADL70639. The invention relates to TSP fragments
CC (80-100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
CC clinical assays for cancer and for generation of antibodies and other
CC binding agents. A method that distinguishes TSP from a TSP fragment or
CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
CC or portion as a target for a binding molecule, e.g. an antibody, to
CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
CC epitope present in TSP but not in the fragment or portion to obtain a
CC quantitation of TSP only; and (3) using the difference between (1) and
CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
CC the TSP fragment or portion is performed in order to detect the presence,
CC or monitor the course, of a disease or condition selected from cancer,

CC renal failure, renal disease, atopic dermatitis, vasculitis, acute
 CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
 CC infarction, liver disease, splenectomy, dermatomyositis, polyarteritis
 CC nodea, systemic lupus erythematosus, lupus erythematosus, Kawasaki
 CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
 CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
 CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition
 CC associated with clotting, a condition associated with platelet
 CC activation, a condition associated with intravascular platelet
 CC -induced thrombocytopenia, disseminated intravascular coagulation,
 CC intravascular coagulation, extravascular coagulation, a condition
 CC associated with endothelial activation, a condition associated with
 CC production and/or release of thrombospondin and/or a thrombospondin
 CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 CC reaction, an aspartame reaction, atopic dermatitis, eczema,
 CC hypersensitivity, scleroderma, conditions associated with plugging of
 CC vessels, a condition associated with a cryofibrinogen, a condition
 CC associated with a cryoglobulin, and a condition associated with an anti-
 CC cardioliipin antibody. The cancer is selected from adenoma.
 CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system,
 CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
 CC biliary system, pancreas, head, neck, endocrine system, reproductive
 CC system (male or female), genitourinary system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.

XX SQ Sequence 240 AA;

Query Match 100.0%; Score 55; DB 8; Length 240;

Best Local Similarity 100.0%; Pred. NO. 0.023;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 190 FQGVQLQNVRFV 200

RESULT 7

ADQ39359

ID ADQ39359 standard; protein; 432 AA.

XX AC ADQ39359;

XX 18-NOV-2004 (first entry)

DE Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW cardiant; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

PR 10-MAR-2003; 2003US-0453135P.

PR 30-APR-2003; 2003US-0466412P.

PR 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakubova O;

XX

DR WPI; 2004-533949/51.
 N-PSDB; ADQ38531.

XX Identifying an individual who has an altered risk for developing
 PT myocardial infarction by detecting a single nucleotide polymorphism in
 PT the individual's nucleic acids.

XX Claim 10; SEQ ID NO 1022; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiac activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC sequence represents the protein of a human myocardial infarction-
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 432 AA;

Query Match 100.0%; Score 55; DB 8; Length 432;

Best Local Similarity 100.0%; Pred. No. 0.044; 0; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 208 FQGVQLQNVRFV 218

RESULT 8

ADQ39357

ID ADQ39357 standard; protein; 432 AA.

XX AC ADQ39357;

XX 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1020.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW cardiant; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

PR 10-MAR-2003; 2003US-0453135P.

PR 30-APR-2003; 2003US-0466412P.

PR 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Takubova O;
XX WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38529.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 10; SEQ ID NO 1020; 145pp; English.
XX
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiac activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 432 AA;
Query Match 100.0%; Score 55; DB 8; Length 432;
Best Local Similarity 100.0%; Pred. NO. 0.044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQLQNVRFV 11
DB 208 FQGVQLQNVRFV 218
RESULT 9
AAU02916
ID AAU02916 standard; protein; 459 AA.
XX
XX AAU02916;
XX
XX 12-SEP-2001 (first entry)
XX
XX Angiotensin converting enzyme (ACEV) splice variant protein #16.
XX
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW noncardioidic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.

XX Homo sapiens.
XX WO200136632-A2.
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-IL000766.
XX
XX 17-NOV-1999; 99IL-00132978.
XX 10-DEC-1999; 99IL-00133455.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX DR N-PSDB; AAS06016.
XX
PT Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 16; 519pp; English.
XX
XX The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding to the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, noncardioidic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 459 AA;
Query Match 100.0%; Score 55; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. NO. 0.047;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQLQNVRFV 11
DB 208 FQGVQLQNVRFV 218
RESULT 10
AAB43602
ID AAB43602 standard; protein; 466 AA.
XX
XX AAB43602;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated protein sequence SEQ ID NO:1047.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiaesthetic; antithrombotic; antithrombotic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.

```

XX OS Homo sapiens.
XX KW WO200055350-A1.
XX PN 21-SEP-2000.
XX PD
XX PF 08-MAR-2000; 2000WO-US005882.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX DR N-PSDB; AAC77811.
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX PS Claim 11; Page 1636-1638; 2352pp; English.
XX KW AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerable; immunomodulator;
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiatic; thrombolytic; coagulant;
XX neotropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention
XX SQ Sequence 466 AA;
      Query Match 100.0%; Score 55; DB 3; Length 466;
      Best Local Similarity 100.0%; Pred. NO. 0.048;
      Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLQNVRFV 11
Db 261 FQGVLQNVRFV 271
      |||||
      |||||

RESULT 11
AAU02915
ID AAU02915 standard; protein; 546 AA.
XX AC AAU02915;
XX DT 12-SEP-2001 (first entry)
XX DE Angiotensin converting enzyme (ACEV) splice variant protein #15.
XX KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX KW platelet-derived endothelial cell growth factor; cardiovascular disease;
XX KW cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
XX KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX KW myocardial infarction; coronary arterial thrombosis; renal disease;
XX KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

```

```

KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX OS Homo sapiens.
XX PN WO200136632-A2.
XX PD 25-MAY-2001.
XX PF 17-NOV-2000; 2000WO-IL000766.
XX PR 17-NOV-1999; 99IL-00132978.
XX PR 10-DEC-1999; 99IL-00133455.
XX PA (COMP-) COMPUGEN LTD.
XX PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX DR N-PSDB; AAS06015.
XX PT Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
XX capable of binding to the variant and to detect anti-variant antibodies.
XX PS Claim 4; Fig 15; 519pp; English.
XX KW The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
XX granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
XX platelet-derived endothelial cell growth factor, cyclin-dependent kinase
XX inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
XX polypeptide receptor 2. The polypeptides and their associated nucleic
XX acids are useful for identification of variant sequences and detection of
XX candidate compounds capable of binding the molecules. The sequences of
XX the invention can be used in the treatment and diagnosis of various
XX disorders including cardiovascular diseases such as arteriosclerosis,
XX myocardial infarction and coronary arterial thrombosis, renal diseases
XX such as diabetic nephropathy, muscular diseases such as hypertrophy,
XX immune disorders such as immune complex nephritis, multiple sclerosis,
XX cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
XX as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
XX SQ Sequence 546 AA;
      Query Match 100.0%; Score 55; DB 4; Length 546;
      Best Local Similarity 100.0%; Pred. No. 0.057;
      Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVLQNVRFV 11
Db 208 FQGVLQNVRFV 218
      |||||
      |||||

RESULT 12
ADN02474
ID ADN02474 standard; protein; 548 AA.
XX AC ADN02474;
XX DT 17-JUN-2004 (first entry)
XX DE TSF polypeptide.
XX KW adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
XX OS Homo sapiens.
XX PN CN1401387-A.
XX PD 12-MAR-2003.

```

XX 21-AUG-2002; 2002CN-00129408.
 XX
 XX 21-AUG-2002; 2002CN-00129408.
 XX
 XX (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
 XX
 XX Han Z, Liu P;
 XX
 XX WPI; 2003-469302/45.
 XX
 XX N-PSDB; ADN02475.
 XX
 XX Tumor suppressing polypeptide TSF and gene therapy vector composition.
 XX
 XX Claim 2; SEQ ID NO 1; 13pp; Chinese.
 XX
 XX The present invention relates to a novel recombinant adenovirus vector
 CC mediated anti-neoplastic composition is prepared through cloning the cDNA
 CC sequence from the human peripheral blood cell by specific primer and
 CC reverse transcription-polymerase chain reaction (RT-PCR) method for
 CC coding TSF polypeptide, constructing in human embryonic kidney cell 293
 CC by ADEasy system, and packaging and expressing the recombinant adenovirus
 CC vector of TSF. It can suppress the growth and transfer of cancer. The
 CC present sequence represents the TSF polypeptide.
 XX
 XX Sequence 548 AA;
 SQ Query Match 100.0%; Score 55; DB 7; Length 548;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
 |||||
 DB 208 FQGVQLQNVRFV 218

RESULT 13
 AAU02914
 ID AAU02914 standard; protein; 555 AA.
 XX
 XX AAU02914;
 XX
 XX 12-SEP-2001 (first entry)
 XX
 XX Angiotensin converting enzyme (ACEV) splice variant protein #14.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonarctoidic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 XX WO200136632-A2.
 XX
 XX 25-MAY-2001.
 XX
 XX 17-NOV-2000; 2000WO-IL000766.
 XX
 XX 17-NOV-1999; 99IL-00132978.
 PR
 PR 10-DEC-1999; 99IL-00133455.
 XX
 XX (COMP-) COMPUGEN LTD.
 XX
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI
 XX WPI; 2001-336004/35.
 XX

DR N-PSDB; AAS06014.
 XX
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 XX Claim 4; Fig 14; 519pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonarctoidic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 XX Sequence 555 AA;
 SQ Query Match 100.0%; Score 55; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 0.058;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
 |||||
 DB 208 FQGVQLQNVRFV 218

RESULT 14
 AAU02913
 ID AAU02913 standard; protein; 731 AA.
 XX
 XX AAU02913;
 XX
 XX 12-SEP-2001 (first entry)
 XX
 XX Angiotensin converting enzyme (ACEV) splice variant protein #13.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonarctoidic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 XX
 OS Homo sapiens.
 XX
 XX WO200136632-A2.
 XX
 XX 25-MAY-2001.
 XX
 XX 17-NOV-2000; 2000WO-IL000766.
 XX
 XX 17-NOV-1999; 99IL-00132978.
 PR
 PR 10-DEC-1999; 99IL-00133455.
 XX
 XX (COMP-) COMPUGEN LTD.
 XX
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI
 XX WPI; 2001-336004/35.
 XX

```

DR N-PSDB; AAS06013.
XX
PT Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 13; 519pp; English.
XX
CC The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonrheumatic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 731 AA;
Query Match 100.0%; Score 55; DB 4; Length 731;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
DB 208 FQGVQLQNVRFV 218

RESULT 15
AAB00042
ID AAB00042 standard; protein; 1152 AA.
XX
AC AAB00042;
XX
DT 08-NOV-2000 (first entry)
XX
DE Human thrombospondon-1 (TSP-1).
XX
KW TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
KW diabetic retinopathy; corneal graft rejection; glaucoma.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 361..416
   /label= Type 1 repeat region
FT Region 417..473
   /label= Type 1 repeat region
FT Region 474..530
   /label= Type 1 repeat region
XX
PN WO200044908-A2.
XX
PD 03-AUG-2000.
XX
PF 01-FEB-2000; 2000WO-US002482.
XX
PR 01-FEB-1999; 99US-0118053P.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Lawler JW;
XX
WPI; 2000-514823/46.
XX

```

```

XX Nucleic acids encoding chimeric proteins such as cartilage oligomeric
PT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
PT inhibiting angiogenesis and treating diseases such as cancer.
XX
PS Disclosure; Fig 1; 40pp; English.
XX
CC New nucleic acids are described which encode a protein comprising the
CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
CC the TGF (transforming growth factor)-beta activation region of human TSP-
CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
CC tumours in mice models. Thus the nucleic acids and proteins may be useful
CC for treating angiogenesis related diseases such as cancer (by reducing
CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
CC for treating human immunodeficiency virus (HIV) infection. Anti-
CC angiogenic therapy has little toxicity, does not require the therapeutic
CC agent to enter tumour cells or cross the blood-brain barrier, controls
CC tumour growth independently of growth of tumour cell heterogeneity, and
CC does not induce drug resistance
XX
SQ Sequence 1152 AA;
Query Match 100.0%; Score 55; DB 3; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
DB 190 FQGVQLQNVRFV 200

Search completed: June 5, 2006, 22:24:59
Job time : 83.5431 secs

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:25:22 ; Search time 12.8017 Seconds
(without alignments)
82.675 Million cell updates/sec

Title: US-10-030-735-30

Perfect score: 55

Sequence: 1 FQGVQLQNVRFV 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	229	S57957	thrombospondin 1 -
2	55	100.0	1170	1 TSHUP1	thrombospondin 1 p
3	55	100.0	1170	2 A40558	thrombospondin 1 p
4	38	69.1	467	2 D84938	H+-transporting tw
5	37	67.3	1172	1 TSHUP2	thrombospondin 2 p
6	37	67.3	1172	2 A42587	thrombospondin 2 p
7	36	65.5	186	2 B75421	probable piliin, ty
8	36	65.5	247	1 A64590	probable 3-oxoacyl
9	36	65.5	247	2 B71923	3-oxoacyl-[acyl-ca
10	36	65.5	1054	2 H69377	reverse gyrase (to
11	35	63.6	304	2 T34271	hypothetical prote
12	35	63.6	498	2 B89832	hypothetical prote
13	35	63.6	740	2 G95153	neuraminidase, pro
14	35	63.6	747	2 AE2929	two component resp
15	35	63.6	759	2 AC0368	probable autotrans
16	35	63.6	783	2 A98353	probable transcrip
17	35	63.6	1178	1 A39804	thrombospondin pre
18	34	61.8	82	2 F96625	hypothetical prote
19	34	61.8	212	2 D81929	probable imidazole
20	34	61.8	295	2 E72462	probable lipocalin
21	34	61.8	459	2 B95171	NADH oxidase Sp146
22	34	61.8	459	2 B98037	NADH oxidase (EC 1
23	34	61.8	715	2 S70397	zona pellucida gly
24	34	61.8	1308	2 T05178	hypothetical prote
25	34	61.8	1829	2 AE1864	hypothetical prote
26	34	61.8	3587	2 I40486	surfactin syntheta
27	33	60.0	93	2 T31048	hypothetical prote
28	33	60.0	298	2 C87403	FdhD protein limpo
29	33	60.0	417	2 H83708	hypothetical prote

30	33	60.0	555	2 T23531	hypothetical prote
31	33	60.0	552	2 T03504	probable DNA topoi
32	33	60.0	750	1 HYRUN	neprilysin (EC 3.4
33	33	60.0	750	1 HYRTN	neprilysin (EC 3.4
34	33	60.0	751	1 HYRBN	neprilysin (EC 3.4
35	33	60.0	876	2 B96693	probable receptor
36	33	60.0	946	2 S71168	Ca2+-transporting
37	33	60.0	1020	2 D86402	protein envelope C
38	33	60.0	1020	2 T51925	Ca2+-transporting
39	33	60.0	1020	2 T51926	Ca2+-transporting
40	33	60.0	1038	2 T02634	rep protein homolo
41	33	60.0	1069	2 C85349	Ca2+-transporting
42	33	60.0	1093	2 T08551	Ca2+-transporting
43	33	60.0	1545	2 T42751	sulfonylurea recep
44	33	60.0	1545	2 T46645	sulfonylurea recep
45	33	60.0	3712	2 S18253	laminin alpha-1 ch

ALIGNMENTS

RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C:Accession: S57957

R:Lafuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

A:Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expres

A:Reference number: S57955

A:Accession: S57957

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-229 <LAF>

A:Cross-references: UNIPROT:Q28194; UNIPARC:UPI000008740A; EMBL:X89511; NID:G899228; PI

C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v

Query Match 100.0%; Score 55; DB 2; Length 229;

Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11

Db 190 FQGVQLQNVRFV 200

RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)

C>Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: A26155; A34274; A30140; A25812; A05172; A42927

R:Lawler, J.; Hyman, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A>Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple

A:Reference number: A26155; MUID:87057617; PMID:2430973

A:Accession: A26155

A:Molecule type: mRNA

A:Residues: 1-1170 <LAW>

A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:G37137; PIDN:

A>Note: parts of this sequence, including the amino end of the mature protein, were det

R:Lahterty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A>Title: Characterization of the promoter region of the human thrombospondin gene. DNA

A:Reference number: A34274; MUID:89291870; PMID:2544587

A:Accession: A34274

A:Molecule type: DNA

A:Residues: 1-166 <LAH>

A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835

R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,

J. Cell Biol. 108, 729-736, 1989

A>Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in t

C;Genetics:
A;Gene: fliI; BU076
C;Superfamily: H(+)-transporting ATP synthase; H+-transporting ATP synthase alpha chain
C;Keywords: hydrolase

Query Match 69.1%; Score 38; DB 2; Length 467;
Best Local Similarity 60.0%; Pred. No. 10;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLNVRV 10
|:|:|:|:|:|
Db 10 FRGLQNLRF 19

RESULT 5
TSHUP2
C;Species: Homo sapiens (man)
C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: A47379; A42173
R;LaBell, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
A;Reference number: A47379; MUID:94010892; PMID:8406456
A;Accession: A47379
A;Molecule type: mRNA
A;Residues: 1-1172 <LAB>
A;Cross-references: UNIPROT:P35442; UNIPARC:UPI00001742C1; GB:M81339
R;LaBell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression o
A;Reference number: A42173; MUID:92217961; PMID:1559694
A;Accession: A42173
A;Molecule type: mRNA
A;Residues: 560-1172 <LA2>
A;Cross-references: UNIPARC:UPI00001742C1; GB:M81339
A;Experimental source: fibroblast
A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
C;Genetics:
A;Gene: GDB:THBS2; TSP2
A;Cross-references: GDB:128789; OMIM:188061
A;Map position: 6q27-6q27
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1172/Product: thrombospondin 2 #status predicted <MAT>
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>
F;928-930/Region: cell attachment (R-G-D) motif
F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predi
F;167-226/Disulfide bonds: #status predicted
F;266,270/Disulfide bonds: interchain #status predicted
F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 67.3%; Score 37; DB 1; Length 1172;
Best Local Similarity 63.6%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 FQGVQLNVRV 11
|:|:|:|:|:|
Db 202 FRGLQNLRF 212

RESULT 6
A42587
thrombospondin 2 precursor - mouse

C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42587; A39851
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: A42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1172 <LAH>
A;Cross-references: UNIPROT:Q03350; UNIPARC:UPI0000029847; GB:L07803; GB:M87275; NID:93
A;Note: sequence extracted from NCBI backbone (NCBIP:81502)
R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A;Reference number: A39851; MUID:91302287; PMID:1712771
A;Accession: A39851
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-873 <BOR>
A;Cross-references: UNIPARC:UPI000016D077; GB:M64866; NID:g201994; PIDN:AAA0432.1; PID
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C;Keywords: calcium binding; glycoprotein
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>

Query Match 67.3%; Score 37; DB 2; Length 1172;
Best Local Similarity 63.6%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLNVRV 11
|:|:|:|:|:|
Db 202 FRGLQNLRF 212

RESULT 7
B75421
probable pilin, type IV - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: B75421
R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-186 <WHI>
A;Cross-references: UNIPROT:Q9RUZ7; UNIPARC:UPI000003D3E03; GB:AE001971; GB:AE000513; NI
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1233
A;Map position: 1

Query Match 65.5%; Score 36; DB 2; Length 186;
Best Local Similarity 87.5%; Pred. No. 9.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLNVR 9
|:|:|:|:|:|
Db 60 QGVLENVR 67

RESULT 8
A64590

probable 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Helicobacter pylori
 C:Species: Helicobacter pylori
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
 C:Accession: A64590
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:197394467; PMID:9252185
 A:Accession: A64590
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-247 <TOM>
 A:Cross-references: UNIPROT:Q25286; UNIPARC:UPI00000D3178; GB:AE000511; NID
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: fatty acid biosynthesis; NAD; oxidoreductase
 F:6-186/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 65.5%; Score 36; DB 1; Length 247;
 Best Local Similarity 50.0%; Pred. No. 13; Mismatches 1; Indels 0; Gaps 0;
 Matches 5; Conservative 4;

QY 1 FQGVQLQNVRF 10
 ::||:|:
 Db 171 YEGALRNIRF 180

RESULT 9
 B71923
 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Helicobacter pylori (strain
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Accession: B71923
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 05-Oct-2004
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: B71923
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-247 <ARN>
 A:Cross-references: UNIPROT:Q9ZL50; UNIPARC:UPI00000D3648; GB:AE001484; GB:AE001439; NID
 A:Experimental source: strain J99
 C:Gene: fabG
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: oxidoreductase
 F:6-186/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 65.5%; Score 36; DB 2; Length 247;
 Best Local Similarity 50.0%; Pred. No. 13; Mismatches 4; Indels 0; Gaps 0;
 Matches 5; Conservative 4;

QY 1 FQGVQLQNVRF 10
 ::||:|:
 Db 171 YEGALRNIRF 180

RESULT 10
 H69377
 reverse gyrase (top-RG) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: H69377
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: H69377
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1054 <KLE>
 A:Cross-references: UNIPROT:Q29238; UNIPARC:UPI0000056EB5; GB:AE001033; GB:AE000782; N1
 C:Keywords: ATP; nucleotide binding; P-loop
 F:78-85/Region: nucleotide-binding motif A (P-loop)
 F:257-262/Region: nucleotide-binding motif B
 F:261-263/Region: DEAD motif #status atypical

Query Match 65.5%; Score 36; DB 2; Length 1054;
 Best Local Similarity 72.7%; Pred. No. 62;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
 |||||
 Db 565 FHGVLVNGRFV 575

RESULT 11
 T34271
 hypothetical protein F46C8.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Dec-2004
 C:Accession: T34271
 R:Wilcox, L.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid F46C8.
 A:Reference number: Z21497
 A:Accession: T34271
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-304 <WIL>
 A:Cross-references: UNIPARC:UPI00001793F8; EMBL:U41624; PIDN:AAA83323.1; CESP:F46C8.8
 C:Genetics:
 A:Gene: CESP:F46C8.8
 A:Introns: 24/1; 56/1; 101/1; 123/1; 149/3; 219/1; 275/1

Query Match 63.6%; Score 35; DB 2; Length 304;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRF 10
 |||||:
 Db 22 FQGVLANAKF 31

RESULT 12
 B9832
 hypothetical protein SA0581 [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C:Accession: B9832
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: B9832
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-438 <KUR>
 A:Cross-references: UNIPROT:Q99VV9; UNIPARC:UPI00000CADE1; GB:EA000018; PID:q13700516;
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: SA0581

Query Match 63.6%; Score 35; DB 2; Length 498;
 Best Local Similarity 54.5%; Pred. No. 44;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
||| ||| :
Db 396 FQGVQNVRFV 406

RESULT 13

G95153
neuraminidase, probable [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: G95153
R;Tetralin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: G95153
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-740 <KUR>
A;Cross-references: UNIPROT:Q97Q99; UNIPARC:UPI00000517D9; GB:AE005672; PIDN:AAK75424.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1326

Query Match 63.6%; Score 35; DB 2; Length 740;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 10
||| ||| :
Db 245 FNGTLENKIF 254

RESULT 14

AE2929
two component response regulator Atu3035 [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AE2929
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, D
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-747 <KUR>
A;Cross-references: UNIPROT:Q8UB11; UNIPARC:UPI0000164787; GB:AE008689; PIDN:AAL43851.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3035
A;Map position: linear chromosome

Query Match 63.6%; Score 35; DB 2; Length 747;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 GVLQNVRFV 11
||| ||| :
Db 42 GILESVRVF 50

RESULT 15

AC0368

probable autotransporter YPO3028 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0368
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0368
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-759 <KUR>
A;Cross-references: UNIPROT:Q8ZC3; UNIPARC:UPI000000DCBBF; GB:AL590842; PIDN:CAC922270.1
C;Genetics:
A;Gene: YPO3028

Query Match 63.6%; Score 35; DB 2; Length 759;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 10
||| ||| :
Db 473 FDGEMQNLRF 482

Search completed: June 5, 2006, 22:45:03
Job time : 12.8017 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:09:41 ; Search time 99.8534 Seconds
(without alignments)
101.901 Million cell updates/sec

Title: US-10-030-735-30
Perfect score: 55
Sequence: 1 FQGVQNVRPV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	229	Q28194_BOVIN	Q28194 bos taurus
2	55	100.0	496	Q7SY84_XENLA	Q7SY84 xenopus lae
3	55	100.0	1170	Q28178_BOVIN	Q28178 bos taurus
4	55	100.0	1170	TSP1_HUMAN	P07996 homo sapien
5	55	100.0	1170	TSP1_MOUSE	P35441 mus musculus
6	55	100.0	1170	Q3TR40_MOUSE	Q3TR40 mus musculus
7	55	100.0	1170	Q718A3_RAT	Q718a3 rattus norv
8	55	100.0	1171	Q80YQ1_MOUSE	Q80YQ1 mus musculus
9	55	100.0	1171	Q8CGB2_MOUSE	Q8CGB2 mus musculus
10	55	100.0	1173	TSP1_XENLA	P35448 xenopus lae
11	55	100.0	1225	Q59E99_HUMAN	Q59E99 homo sapien
12	50	90.9	1090	Q5SPG5_BRARE	Q5SPG5 brachydanio
13	50	90.9	1193	Q4S758_TETNG	Q4S758 tetraodon n
14	49	89.1	249	Q5U903_PIG	Q5U903 sus scrofa
15	49	89.1	1171	Q4RLR5_TETNG	Q4RLR5 tetraodon n
16	45	81.8	1034	Q4RQ74_TETNG	Q4RQ74 tetraodon n
17	42	76.4	704	Q4P665_USTWA	Q4P665 ustilago ma
18	42	76.4	727	Q8PRY3_METWA	Q8PRY3 methanosarc
19	42	76.4	730	Q467V1_METFA	Q467V1 methanosarc
20	42	76.4	733	Q8TLX6_METHAC	Q8TLX6 methanosarc
21	42	76.4	1549	Q3CUK9_THEAT	Q3CUK9 thermoaer
22	41	74.5	713	Q3FIU8_9BURK	Q3FIU8 burkholderi
23	41	74.5	713	Q44XL2_9BURK	Q44XL2 burkholderi
24	41	74.5	713	Q4LLM8_9BURK	Q4LLM8 burkholderi
25	41	74.5	1168	Q5VH52_CIOIN	Q5VH52 ciona intes
26	40	72.7	695	Q2LZ43_DROPS	Q2LZ43 drosophila
27	39	70.9	189	Q3LIL6_NEIGI	Q3LIL6 thiomicrosp
28	39	70.9	212	Q5FA21_NEIGI	Q5FA21 neisseria g
29	39	70.9	494	Q5AVQ8_EMENI	Q5AVQ8 aspergillus
30	39	70.9	595	Q86NR6_DROME	Q86NR6 drosophila
31	39	70.9	721	Q9VTH0_DROME	Q9VTH0 drosophila

32	39	70.9	1034	2	Q4DWH6_TRYCR	Q4DWH6 trypanosoma
33	39	70.9	1034	2	Q7Z2B9_TRYCR	Q7Z2B9 trypanosoma
34	39	70.9	1172	2	Q6FXF4_CANGA	Q6FXF4 candida gla
35	39	70.9	1367	2	Q2TYI8_ASPLI	Q2TYI8 aspergillus
36	38	69.1	182	2	Q6C422_YARLI	Q6C422 yarrowia li
37	38	69.1	467	1	FLII_BUCAI	P57178 buchera ap
38	38	69.1	3242	2	Q859P9_BPN4	Q859P9 bacterioph
39	37	67.3	261	2	Q3C7L8_9CLOT	Q3C7L8 alkaliphilu
40	37	67.3	438	2	Q41VB6_DESHA	Q41VB6 desulfitoba
41	37	67.3	750	2	Q8BVV4_MOUSE	Q8BVV4 mus musculu
42	37	67.3	1172	1	TSP2_HUMAN	P35442 homo sapien
43	37	67.3	1172	1	TSP2_MOUSE	Q03350 mus musculu
44	37	67.3	1172	2	Q5RI52_HUMAN	Q5RI52 homo sapien
45	37	67.3	1172	2	Q7TWT3_MOUSE	Q7TWT3 mus musculu

ALIGNMENTS

RESULT 1
Q28194_BOVIN PRELIMINARY; PRT; 229 AA.
AC Q28194;
DT 01-NOV-1996, integrated into UniProtKB/TREMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Thrombospondin-1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96331130; PubMed=8698834;
RX DOI=10.1002/(SICI)1097-4652(199604)167:1<164::AID-JCP19>3.3.CO;2-O;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
secreted protein/thrombospondin-2 expression by adrenocorticotrophic
hormone in adrenocortical cells.";
J. Cell. Physiol. 167:164-172(1996).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; X89511; CAA61682.1; -; mRNA.
DR PIR; S57957; S57957.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR013320; ConA like subgrp.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR SMART; SM00210; TSPN; 1.
FT NON_TER 1
FT NON_TER 229
SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;
Query Match 100.0%; Score 55; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQNVRPV 11
DB 190 FQGVQNVRPV 200
RESULT 2
Q7SY84_XENLA PRELIMINARY; PRT; 496 AA.
AC Q7SY84;
DT 01-OCT-2003, integrated into UniProtKB/TREMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 12.

DE MGC64438 protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]_TaxID=8355;
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Liu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Whole;
 RX Klein S., Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; BC054970; AAH54970.1; -; mRNA.
 DR GO; GO:0005198; F.structural molecule activity; IEA.
 DR GO; GO:0007155; P.cell adhesion; IEA.
 DR InterPro; IPR013320; ConA like subgrp.
 DR InterPro; IPR003129; Laminin_G_TSP_N.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP 1.
 DR InterPro; IPR001007; VWFC.
 DR Pfam; PF00090; TSP_1; 2.
 DR Pfam; PF00093; VWFC; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00209; TSPN; 1.
 DR SMART; SM00214; VWFC; 1.
 DR PROSITE; PS0092; TSP1; 2.
 DR PROSITE; PS01208; VWFC 1; UNKNOWN_1.
 DR PROSITE; PS0184; VWFC 2; 1.
 DR PROSITE; PS0184; VWFC 2; 1.
 DR SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7EF51B CRC64;

Query Match 100.00%; Score 55; DB 2; Length 496;

Best Local Similarity 100.00%; Pred. No. 0.047; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGLQNVRFV 11

|||||

DB 214 FQGLQNVRFV 224
 RESULT 3
 TSP1_BOVIN
 ID TSP1_BOVIN STANDARD; PRT; 1170 AA.
 AC Q28178; Q28179;
 DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
 DT 01-DEC-2000, sequence version 2.
 DT 07-MAR-2006, entry version 56.
 DE Thrombospondin-1 precursor.
 GN Name=THBS1; Synonyms=TSP-1, TSP1;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RC STRAIN=Holstein; TISSUE=Tooth;
 RX MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
 RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
 RA Inoue H.;
 RT "cDNA cloning of bovine thrombospondin 1 and its expression in
 RT odontoblasts and predentin.";
 RL Biochim. Biophys. Acta 1382:17-22(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 1-18 AND 710-1170.
 RC TISSUE=Aortic endothelium;
 RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
 RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
 RT TGF-beta.";
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
 CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
 CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
 CC V/beta-3 and alpha-11b/beta-3. May play a role in dentinogenesis
 CC and/or maintenance of dentin and dental pulp.
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.
 CC -!- TISSUE SPECIFICITY: Odontoblasts.
 CC -!- SIMILARITY: Belongs to the thrombospondin family.
 CC -!- SIMILARITY: Contains 3 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
 CC -!- SIMILARITY: Contains 1 VWFC domain.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; AB005287; BAA21115.1; -; mRNA.
 DR EMBL; X87618; CAA60950.1; -; mRNA.
 DR EMBL; X87619; CAA60951.1; -; mRNA.
 DR PIR; S55501; S55501.
 DR HSP; P07996; 1LSL.
 DR SMR; Q28178; 549-1169.
 DR GlycoSuiteDB; Q28178; --
 DR InterPro; IPR013320; ConA_like_subgrp.
 DR InterPro; IPR006210; EGF_3.
 DR InterPro; IPR000742; EGF_3.
 DR InterPro; IPR001881; EGF_Ca_bd.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR013032; EGF_like_reg.
 DR InterPro; IPR003129; Laminin_G_TSP_N.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP 1.
 DR InterPro; IPR003367; tsp 3.
 DR InterPro; IPR008859; TSP_C.
 DR InterPro; IPR001007; VWFC.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00090; TSP_1; 3.
 DR Pfam; PF02412; TSP_3; 13.

```
DR PFAM: PF05735; TSP_C; 1.
DR PFAM: PF00093; VWC; 1.
DR PRINTS: PRO1705; TSP1REPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00026; TSP1; 3.
DR PROSITE: PS00092; TSP1; 3.
DR PROSITE: PS01208; VWC; 1.
DR PROSITE: PS0184; VWC; 2; 1.
DR CALCIUM: Call adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT REGION 19 232
FT MOTIF 926 928
FT CARBOHYD 248 248
FT CARBOHYD 360 360
FT CARBOHYD 708 708
FT CARBOHYD 1067 1067
FT CARBOHYD 1085 1085
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 321 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489
FT DISULFID 462 474
FT DISULFID 504 541
FT DISULFID 508 546
FT DISULFID 519 531
FT DISULFID 551 562
FT DISULFID 556 572
FT DISULFID 575 586
FT DISULFID 592 608
FT DISULFID 599 617
FT DISULFID 620 644
FT DISULFID 650 663
FT DISULFID 657 676
FT DISULFID 678 689
FT DISULFID 705 713
FT DISULFID 718 738
FT DISULFID 754 774
FT DISULFID 777 797
FT DISULFID 813 833
FT DISULFID 836 856
FT DISULFID 874 894
FT DISULFID 910 930
FT DISULFID 946 1167
FT CONFLICT 805 805
FT SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Pfam: PF05735; TSP_C; 1.
Pfam: PF00093; VWC; 1.
PRINTS: PRO1705; TSP1REPEAT.
SMART: SM00181; EGF; 3.
SMART: SM00209; TSP1; 3.
SMART: SM00210; TSPN; 1.
SMART: SM00214; VWC; 1.
PROSITE: PS00022; EGF_1; FALSE_NEG.
PROSITE: PS01186; EGF_2; 1.
PROSITE: PS00026; TSP1; 3.
PROSITE: PS00092; TSP1; 3.
PROSITE: PS01208; VWC; 1.
PROSITE: PS0184; VWC; 2; 1.
CALCIUM: Call adhesion; EGF-like domain; Glycoprotein;
Heparin-binding; Repeat; Signal.
SIGNAL 1 18
CHAIN 19 1170
DOMAIN 24 221
DOMAIN 316 373
DOMAIN 379 429
DOMAIN 435 490
DOMAIN 492 547
DOMAIN 549 587
DOMAIN 588 645
DOMAIN 646 690
DOMAIN 723 758
DOMAIN 759 781
DOMAIN 782 817
DOMAIN 818 840
DOMAIN 841 878
DOMAIN 879 914
DOMAIN 915 950
DOMAIN 951 1170
REGION 19 232
MOTIF 926 928
CARBOHYD 248 248
CARBOHYD 360 360
CARBOHYD 708 708
CARBOHYD 1067 1067
CARBOHYD 1085 1085
DISULFID 270 270
DISULFID 274 274
DISULFID 321 423
DISULFID 395 428
DISULFID 406 413
DISULFID 447 484
DISULFID 451 489
DISULFID 462 474
DISULFID 504 541
DISULFID 508 546
DISULFID 519 531
DISULFID 551 562
DISULFID 556 572
DISULFID 575 586
DISULFID 592 608
DISULFID 599 617
DISULFID 620 644
DISULFID 650 663
DISULFID 657 676
DISULFID 678 689
DISULFID 705 713
DISULFID 718 738
DISULFID 754 774
DISULFID 777 797
DISULFID 813 833
DISULFID 836 856
DISULFID 874 894
DISULFID 910 930
DISULFID 946 1167
CONFLICT 805 805
SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 100.0%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

QY 1 FQGVQLQVRV 11
DB 208 FQGVQLQVRV 218

RESULT 4
TSP1_HUMAN
ID TSP1_HUMAN STANDARD; PRT; 1170 AA.
AC P07996; Q15667;
DT 01-AUG-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1988, sequence version 1.
DT 07-MAR-2006, entry version 78.
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endothelial cell;
RX MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologues with several different
RT proteins."
RL J. Cell Biol. 103:1635-1648 (1986).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89139590; PubMed=2918029; DOI=10.1083/jcb.108.2.729;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region."
RL J. Cell Biol. 108:729-736 (1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-397.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins."
RL Biochemistry 25:8418-8425 (1986).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin."
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453 (1986).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription."
RL J. Biol. Chem. 264:11222-11227 (1989).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA."
RL Submitted (DEC 1992) to the EMBL/GenBank/DBD databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RX TISSUE=Platelet; PubMed=21125860; DOI=10.1074/jbc.M008073200;
```

RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RL module.";
RN J. Biol. Chem. 276:6485-6498 (2001).
RP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=2238361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339 (2002).
RN [9]
RP CARBOHYDRATE-LINKAGE SITES ASN-248 AND ASN-1067.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Grisenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
hydrazone chemistry, and mass spectrometry.";
RL J. Proteome Res. 4:2070-2080 (2005).
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3.
CC -!- SURUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; M25631; AAA36741.1; -; mRNA.
DR EMBL; X04665; CAA28370.1; -; mRNA.
DR EMBL; X14787; CAA32889.1; -; mRNA.
DR EMBL; M14226; AAA61237.1; ALT SEQ; mRNA.
DR EMBL; J04935; AAA61178.1; -; Genomic_DNA.
DR EMBL; M99425; AAB59366.1; -; mRNA.
DR PIR; A26155; TSHUP1.
DR PDB; 1LSL; X-ray; A=434-546.
DR PDB; 1UX6; X-ray; A=834-1170.
DR PDB; 1Z78; X-ray; A=19-233.
DR PDB; 1Z44; X-ray; A=19-257.
DR PDB; 2ERF; X-ray; A=25-233.
DR SMR; P07996; 549-1169.
DR GlycoSuiteDB; P07996; -.
DR OGP; P07996; -.
DR Ensembl; ENSG00000137801; Homo sapiens.
DR HGNC; HGNC:11785; THBS1.
DR MIM; 188060; gene.
DR Reactome; P07996; -.
DR GO; GO:0005576; C:extracellular region; NAS.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0004871; F:signal transducer activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR013032; EGF like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWF C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.

DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWFC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS00184; VWFC_2; 1.
KW 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT /FTid=PRO_0000035842.
FT TSP N-terminal.
FT VWFC.
FT TSP type-1 1.
FT TSP type-1 2.
FT TSP type-1 3.
FT EGF-like 1.
FT EGF-like 2; calcium-binding (Potential).
FT EGF-like 3.
FT TSP type-3 1.
FT TSP type-3 2.
FT TSP type-3 3.
FT TSP type-3 4.
FT TSP type-3 5.
FT TSP type-3 6.
FT TSP type-3 7.
FT TSP C-terminal.
FT Heparin-binding (Potential).
FT Cell attachment site (Potential).
FT N-linked (GlcNAc...); (Potential).
FT N-linked (GlcNAc...); (Potential).
FT C-linked (Man).
FT /FTid=CAR_000205.
FT O-linked (Fuc...);
FT /FTid=CAR_000206.
FT C-linked (Man).
FT /FTid=CAR_000207.
FT C-linked (Man).
FT /FTid=CAR_000208.
FT O-linked (Fuc...);
FT /FTid=CAR_000209.
FT C-linked (Man).
FT /FTid=CAR_000210.
FT O-linked (Fuc...);
FT /FTid=CAR_000211.
FT N-linked (GlcNAc...); (Potential).
FT N-linked (GlcNAc...);
FT Interchain (Probable).
FT Interchain (Probable).

Query Match 100.0%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
|||||

Db 208 FQGVQNVRFV 218

RESULT 5

TS2P_MOUSE

ID TS2P_MOUSE STANDARD; PRT; 1170 AA.

AC P35411;

DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.

DT 01-JUN-1994, sequence version 1.

DT 07-MAR-2006, entry version 57.

DE Thrombospondin-1 precursor.

DE Thrombospondin-1 precursor.

GN Name=Thbs1; Synonyms=Tsp1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridea; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RP [1]

RP NUCLEOTIDE SEQUENCE.

RP MEDLINE=92128941; PubMed=1774063;

RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,

RA Jenkins N.A.;

RA "Characterization of the murine thrombospondin gene.";

RL Genomics 11:587-600(1991).

RP [2]

RP NUCLEOTIDE SEQUENCE.

RP MEDLINE=92147683; PubMed=1371115;

RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,

RA Dixit V.M.;

RA "Characterization of mouse thrombospondin 2 sequence and expression

RT during cell growth and development.";

RL J. Biol. Chem. 267:3274-3281(1992).

RP [3]

RP NUCLEOTIDE SEQUENCE OF 1-490.

RP MEDLINE=90375545; PubMed=2398070;

RA Bornstein P., Alfi D., Devaryalu S., Framson P., Li P.;

RA "Characterization of the mouse thrombospondin gene and evaluation of

RT the role of the first intron in human gene expression.";

RL J. Biol. Chem. 265:16691-16698(1990).

RP [4]

RP PROTEIN SEQUENCE OF 19-37.

RP PubMed=8654563; DOI=10.1016/0014-5793(96)00460-7;

RA Chen H., Aeschlimann D., Nowlin J., Mosher D.F.;

RA "Expression and initial characterization of recombinant mouse

RT thrombospondin 1 and thrombospondin 3.";

RL FEBS Lett. 387:36-41(1996).

CC [1] FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and

CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,

CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-

CC beta-3 and alpha-IIB/beta-3.

CC [2] SUBUNIT: Homotrimer; disulfide-linked.

CC [3] SIMILARITY: Belongs to the thrombospondin family.

CC [4] SIMILARITY: Contains 3 EGF-like domains.

CC [5] SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.

CC [6] SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

CC [7] SIMILARITY: Contains 3 TSP type-1 domains.

CC [8] SIMILARITY: Contains 7 TSP type-3 domains.

CC [9] SIMILARITY: Contains 1 VWFC domain.

CC [10] SIMILARITY: Contains 1 VWFC domain.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC [11] SIMILARITY: Contains 1 VWFC domain.

CC [12] SIMILARITY: Contains 1 VWFC domain.

CC [13] SIMILARITY: Contains 1 VWFC domain.

CC [14] SIMILARITY: Contains 1 VWFC domain.

CC [15] SIMILARITY: Contains 1 VWFC domain.

CC [16] SIMILARITY: Contains 1 VWFC domain.

CC [17] SIMILARITY: Contains 1 VWFC domain.

CC [18] SIMILARITY: Contains 1 VWFC domain.

CC [19] SIMILARITY: Contains 1 VWFC domain.

CC [20] SIMILARITY: Contains 1 VWFC domain.

CC [21] SIMILARITY: Contains 1 VWFC domain.

CC [22] SIMILARITY: Contains 1 VWFC domain.

CC [23] SIMILARITY: Contains 1 VWFC domain.

CC [24] SIMILARITY: Contains 1 VWFC domain.

CC [25] SIMILARITY: Contains 1 VWFC domain.

CC [26] SIMILARITY: Contains 1 VWFC domain.

CC [27] SIMILARITY: Contains 1 VWFC domain.

CC [28] SIMILARITY: Contains 1 VWFC domain.

CC [29] SIMILARITY: Contains 1 VWFC domain.

CC [30] SIMILARITY: Contains 1 VWFC domain.

CC [31] SIMILARITY: Contains 1 VWFC domain.

CC [32] SIMILARITY: Contains 1 VWFC domain.

CC [33] SIMILARITY: Contains 1 VWFC domain.

CC [34] SIMILARITY: Contains 1 VWFC domain.

CC [35] SIMILARITY: Contains 1 VWFC domain.

CC [36] SIMILARITY: Contains 1 VWFC domain.

CC [37] SIMILARITY: Contains 1 VWFC domain.

CC [38] SIMILARITY: Contains 1 VWFC domain.

CC [39] SIMILARITY: Contains 1 VWFC domain.

CC [40] SIMILARITY: Contains 1 VWFC domain.

CC [41] SIMILARITY: Contains 1 VWFC domain.

CC [42] SIMILARITY: Contains 1 VWFC domain.

CC [43] SIMILARITY: Contains 1 VWFC domain.

CC [44] SIMILARITY: Contains 1 VWFC domain.

CC [45] SIMILARITY: Contains 1 VWFC domain.

CC [46] SIMILARITY: Contains 1 VWFC domain.

CC [47] SIMILARITY: Contains 1 VWFC domain.

CC [48] SIMILARITY: Contains 1 VWFC domain.

CC [49] SIMILARITY: Contains 1 VWFC domain.

CC [50] SIMILARITY: Contains 1 VWFC domain.

CC [51] SIMILARITY: Contains 1 VWFC domain.

CC [52] SIMILARITY: Contains 1 VWFC domain.

CC [53] SIMILARITY: Contains 1 VWFC domain.

CC [54] SIMILARITY: Contains 1 VWFC domain.

CC [55] SIMILARITY: Contains 1 VWFC domain.

CC [56] SIMILARITY: Contains 1 VWFC domain.

CC [57] SIMILARITY: Contains 1 VWFC domain.

CC [58] SIMILARITY: Contains 1 VWFC domain.

CC [59] SIMILARITY: Contains 1 VWFC domain.

CC [60] SIMILARITY: Contains 1 VWFC domain.

CC [61] SIMILARITY: Contains 1 VWFC domain.

CC [62] SIMILARITY: Contains 1 VWFC domain.

CC [63] SIMILARITY: Contains 1 VWFC domain.

CC [64] SIMILARITY: Contains 1 VWFC domain.

CC [65] SIMILARITY: Contains 1 VWFC domain.

CC [66] SIMILARITY: Contains 1 VWFC domain.

CC [67] SIMILARITY: Contains 1 VWFC domain.

CC [68] SIMILARITY: Contains 1 VWFC domain.

CC [69] SIMILARITY: Contains 1 VWFC domain.

CC [70] SIMILARITY: Contains 1 VWFC domain.

CC [71] SIMILARITY: Contains 1 VWFC domain.

CC [72] SIMILARITY: Contains 1 VWFC domain.

CC [73] SIMILARITY: Contains 1 VWFC domain.

CC [74] SIMILARITY: Contains 1 VWFC domain.

CC [75] SIMILARITY: Contains 1 VWFC domain.

CC [76] SIMILARITY: Contains 1 VWFC domain.

CC [77] SIMILARITY: Contains 1 VWFC domain.

CC [78] SIMILARITY: Contains 1 VWFC domain.

CC [79] SIMILARITY: Contains 1 VWFC domain.

CC [80] SIMILARITY: Contains 1 VWFC domain.

CC [81] SIMILARITY: Contains 1 VWFC domain.

CC [82] SIMILARITY: Contains 1 VWFC domain.

CC [83] SIMILARITY: Contains 1 VWFC domain.

CC [84] SIMILARITY: Contains 1 VWFC domain.

CC [85] SIMILARITY: Contains 1 VWFC domain.

CC [86] SIMILARITY: Contains 1 VWFC domain.

CC [87] SIMILARITY: Contains 1 VWFC domain.

CC [88] SIMILARITY: Contains 1 VWFC domain.

CC [89] SIMILARITY: Contains 1 VWFC domain.

CC [90] SIMILARITY: Contains 1 VWFC domain.

CC [91] SIMILARITY: Contains 1 VWFC domain.

CC [92] SIMILARITY: Contains 1 VWFC domain.

CC [93] SIMILARITY: Contains 1 VWFC domain.

CC [94] SIMILARITY: Contains 1 VWFC domain.

CC [95] SIMILARITY: Contains 1 VWFC domain.

CC [96] SIMILARITY: Contains 1 VWFC domain.

CC [97] SIMILARITY: Contains 1 VWFC domain.

CC [98] SIMILARITY: Contains 1 VWFC domain.

CC [99] SIMILARITY: Contains 1 VWFC domain.

CC [100] SIMILARITY: Contains 1 VWFC domain.

CC [101] SIMILARITY: Contains 1 VWFC domain.

CC [102] SIMILARITY: Contains 1 VWFC domain.

CC [103] SIMILARITY: Contains 1 VWFC domain.

CC [104] SIMILARITY: Contains 1 VWFC domain.

CC [105] SIMILARITY: Contains 1 VWFC domain.

CC [106] SIMILARITY: Contains 1 VWFC domain.

CC [107] SIMILARITY: Contains 1 VWFC domain.

CC [108] SIMILARITY: Contains 1 VWFC domain.

CC [109] SIMILARITY: Contains 1 VWFC domain.

CC [110] SIMILARITY: Contains 1 VWFC domain.

CC [111] SIMILARITY: Contains 1 VWFC domain.

CC [112] SIMILARITY: Contains 1 VWFC domain.

CC [113] SIMILARITY: Contains 1 VWFC domain.

CC [114] SIMILARITY: Contains 1 VWFC domain.

CC [115] SIMILARITY: Contains 1 VWFC domain.

CC [116] SIMILARITY: Contains 1 VWFC domain.

CC [117] SIMILARITY: Contains 1 VWFC domain.

CC [118] SIMILARITY: Contains 1 VWFC domain.

CC [119] SIMILARITY: Contains 1 VWFC domain.

CC [120] SIMILARITY: Contains 1 VWFC domain.

CC [121] SIMILARITY: Contains 1 VWFC domain.

CC [122] SIMILARITY: Contains 1 VWFC domain.

CC [123] SIMILARITY: Contains 1 VWFC domain.

CC [124] SIMILARITY: Contains 1 VWFC domain.

CC [125] SIMILARITY: Contains 1 VWFC domain.

CC [126] SIMILARITY: Contains 1 VWFC domain.

CC [127] SIMILARITY: Contains 1 VWFC domain.

CC [128] SIMILARITY: Contains 1 VWFC domain.

CC [129] SIMILARITY: Contains 1 VWFC domain.

CC [130] SIMILARITY: Contains 1 VWFC domain.

CC [131] SIMILARITY: Contains 1 VWFC domain.

CC [132] SIMILARITY: Contains 1 VWFC domain.

CC [133] SIMILARITY: Contains 1 VWFC domain.

CC [134] SIMILARITY: Contains 1 VWFC domain.

CC [135] SIMILARITY: Contains 1 VWFC domain.

CC [136] SIMILARITY: Contains 1 VWFC domain.

CC [137] SIMILARITY: Contains 1 VWFC domain.

CC [138] SIMILARITY: Contains 1 VWFC domain.

CC [139] SIMILARITY: Contains 1 VWFC domain.

CC [140] SIMILARITY: Contains 1 VWFC domain.

CC [141] SIMILARITY: Contains 1 VWFC domain.

CC [142] SIMILARITY: Contains 1 VWFC domain.

CC [143] SIMILARITY: Contains 1 VWFC domain.

CC [144] SIMILARITY: Contains 1 VWFC domain.

CC [145] SIMILARITY: Contains 1 VWFC domain.

CC [146] SIMILARITY: Contains 1 VWFC domain

```

FT CARBOHYD 1067 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 Interchain (Probable).
FT DISULFID 274 Interchain (Probable).
FT DISULFID 351 By similarity.
FT DISULFID 395 By similarity.
FT DISULFID 406 By similarity.
FT DISULFID 447 By similarity.
FT DISULFID 451 By similarity.
FT DISULFID 462 By similarity.
FT DISULFID 504 By similarity.
FT DISULFID 508 By similarity.
FT DISULFID 519 By similarity.
FT DISULFID 551 By similarity.
FT DISULFID 556 By similarity.
FT DISULFID 575 By similarity.
FT DISULFID 592 By similarity.
FT DISULFID 599 By similarity.
FT DISULFID 620 By similarity.
FT DISULFID 650 By similarity.
FT DISULFID 657 By similarity.
FT DISULFID 678 By similarity.
FT DISULFID 705 By similarity.
FT DISULFID 718 By similarity.
FT DISULFID 754 By similarity.
FT DISULFID 777 By similarity.
FT DISULFID 813 By similarity.
FT DISULFID 836 By similarity.
FT DISULFID 874 By similarity.
FT DISULFID 910 By similarity.
FT DISULFID 946 By similarity.
FT CONFLICT 1025 F -> L (in Ref. 2).
SQ SEQUENCE 1170 AA; 129647 MW; 0443E493615E7F06 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. NO. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
DB 208 FQGVQNVRFV 218

RESULT 6
Q3TR40 MOUSE
ID Q3TR40_MOUSE PRELIMINARY; PRT; 1170 AA.
AC Q3TR40;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 5.
DE Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DE clone:A530055N06 product:thrombospondin 1, full insert sequence.
GN Name=Tbbs1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambeesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,

```

```

RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioloni M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Putaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiu K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=223546851; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteu G., Pesole G.,
RA Petrovsky N., Pillai K., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yamanishi A., Yoshino M., Waterston R., Iander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```

RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Iehi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AK163092; BAE37190.1; -; mRNA.
CC MGI: MGI:98737; Thbs1.
CC GO: GO:0005615; C:extracellular space; IDA.
CC GO: GO:0005615; C:extracellular space; RCA.
CC GO: GO:0016525; P:negative regulation of angiogenesis; IDA.
CC InterPro: IPR006210; EGF.
CC InterPro: IPR000742; EGF 3.
CC InterPro: IPR001881; EGF_Ca_bd.
CC InterPro: IPR006209; EGF_like.
CC InterPro: IPR013032; EGF_like_reg.
CC -----
CC EMBL: AF309630; AAQ14549.1; -; mRNA.
CC SMR: Q71SA3; 834-1169.
CC DR GO: GO:0005576; C:extracellular region; IEA.
CC DR GO: GO:0005509; F:calcium ion binding; IEA.
CC DR GO: GO:0003198; F:structural molecule activity; IEA.
CC DR GO: GO:0007155; P:cell adhesion; IEA.
CC InterPro: IPR013320; ConA_like_subgrp.
CC InterPro: IPR006210; EGF.
CC InterPro: IPR000742; EGF 3.
CC InterPro: IPR001881; EGF_Ca_bd.
CC InterPro: IPR006209; EGF_like.
CC InterPro: IPR013032; EGF_like_reg.
CC InterPro: IPR003129; Laminin_G_TSP_N.
CC InterPro: IPR000884; TSPI.
CC InterPro: IPR008085; TSP 1.
CC InterPro: IPR003387; tsp_3.
CC InterPro: IPR008859; TSP_C.

DR InterPro: IPR003129; Laminin_G_TSP_N.
DR InterPro: IPR000884; TSPI.
DR InterPro: IPR008085; TSP 1.
DR InterPro: IPR003387; tsp_3.
DR InterPro: IPR008859; TSP_C.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00090; TSP 1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VWC; 1.
DR PRINTS: PR01705; TSPIREPEAT.
DR SMART: SM00181; EGF; 3.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01186; EGF 2; UNKNOWN_1.
DR PROSITE: PS00026; EGF 3; 2.
DR PROSITE: PS50092; TSPI; 3.
DR PROSITE: PS01208; WFC 1; 1.
DR PROSITE: PS50184; WFC_2; 1.
SQ SEQUENCE 1170 AA; 129619 MW; 55BF04A20D91B194 CRC64;
Query Match 100.0%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;
QY 1 FQGVQLQNVRFV 11
DB 208 FQGVQLQNVRFV 218
AC Q71SA3 PRELIMINARY; PRT; 1170 AA.
ID Q71SA3 RAT
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Thrombospondin 1.
GN Names=Tsp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RA Iwabu A., Hirohata S., Kusachi S., Nakamura K., Murakami T.,
RA Ninomiya Y., Tsuji T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AF309630; AAQ14549.1; -; mRNA.
CC SMR: Q71SA3; 834-1169.
CC DR GO: GO:0005576; C:extracellular region; IEA.
CC DR GO: GO:0005509; F:calcium ion binding; IEA.
CC DR GO: GO:0003198; F:structural molecule activity; IEA.
CC DR GO: GO:0007155; P:cell adhesion; IEA.
CC InterPro: IPR013320; ConA_like_subgrp.
CC InterPro: IPR006210; EGF.
CC InterPro: IPR000742; EGF 3.
CC InterPro: IPR001881; EGF_Ca_bd.
CC InterPro: IPR006209; EGF_like.
CC InterPro: IPR013032; EGF_like_reg.
CC InterPro: IPR003129; Laminin_G_TSP_N.
CC InterPro: IPR000884; TSPI.
CC InterPro: IPR008085; TSP 1.
CC InterPro: IPR003387; tsp_3.
CC InterPro: IPR008859; TSP_C.

```

DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; SM01705; TSP1REPEAT.
DR SMART; SM00181; EGF_3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS0026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C_1; UNKNOWN_1.
DR PROSITE; PS0184; VWF_C_2; 1.
SQ SEQUENCE 1170 AA; 129671 MW; 6F38D3DCE733060F CRC64;

Query Match 100.0%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQVQLQNVRFV 11
Db 208 FQVQLQNVRFV 218
|||||

RESULT 8
Q80YQ1 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q80YQ1.
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Thrombospondin 1.
GN Name=Thbs1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

```

CC EMBL; BC050917; AAH50917.1; -; mRNA.
DR HSSP; P07996; 1LSI.
DR SMR; Q80YQ1; 835-1170.
DR Ensembl; ENSMUSG00000040152; Mus musculus.
DR MGI; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; RCA.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF Ca bd.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR013032; EGF-like reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF_3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWF_C_1; UNKNOWN_1.
DR PROSITE; PS0184; VWF_C_2; 1.
SQ SEQUENCE 1171 AA; 129690 MW; 12E077B50C64E2D3 CRC64;

Query Match 100.0%; Score 55; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0;

QY 1 FQVQLQNVRFV 11
Db 208 FQVQLQNVRFV 218
|||||

RESULT 9
Q8CGB2 MOUSE PRELIMINARY; PRT; 1171 AA.
AC Q8CGB2.
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Thrombospondin 1 (Mammary gland RCB-0527 JY9-MC(B) cDNA, RIKEN full-length enriched library, clone:G930018021 product:thrombospondin 1, full insert sequence).
DE full insert sequence.
DE Name=Thbs1.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Mammary tumor metastasized to lung. Tumor
arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Trichmond J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Groomwood J.J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=CZECH I1; TISSUE=Mammary tumor metastasized to lung. Tumor
 RC arose spontaneously;
 RC Straussberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC PubMed=16141072; DOI=10.1126/science.1112014;
 RX Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Banal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
 RA Tamajima K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessky C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami N., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 RN [5]
 RN NUCLEOTIDE SEQUENCE.

RC TISSUE=Mammary gland;
 RX PubMed=16141072; DOI=10.1126/science.1112009;
 RX RIKEN Genome Exploration Research Group, and Genome Science Group
 RG (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense Transcription in the Mammalian Transcriptome.";
 RL Science 309:1564-1566(2005).
 RN [6]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kiyosawa H.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bragati T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Dalla E., Dragani T.A., Pletcher C., Godzik A., Gough J.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kongsaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayateu N.,
 RA Hironaka-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [7]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohteki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [8]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Mammary gland;
 RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RX Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [9]

```

RP NUCLEOTIDE SEQUENCE.
RX TISSUE=Mammary gland;
RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (KISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[10]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC042422; AAH42422.1; -; mRNA.
DR EMBL; AK145202; BAE26293.1; -; mRNA.
DR HSSP; P07996; 1LSL.

Query Match 100.0%; Score 55; DB 2; Length 1171;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 FQGVQLNVRVF 11
Db 208 FQGVQLNVRVF 218

RESULT 10
TSPI_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-1994, sequence version 1.
DT 07-MAR-2006, entry version 54.
DE Thrombospondin-1 precursor.
GN Name=thbs1; Synonyms=tspl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RA "Cloning, characterization and expression of thrombospondin-1 in
RT Xenopus laevis embryos.";
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-1ib/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.

```

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; L04278; -; NOT_ANNOTATED_CDS; mRNA.
DR HSSP; P07996; 1LSL.
DR SMR; P35448; 552-1172.
DR InterPro; IPR013320; ConA_like_subgrp.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR00742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS00184; VWFC_2; 1.
KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1173 Thrombospondin-1.
FT DOMAIN 23 224 /FTID=PRO 0000035844.
FT DOMAIN 319 376 TSP N-terminal.
FT DOMAIN 382 432 TSP type-1 1.
FT DOMAIN 438 493 TSP type-1 2.
FT DOMAIN 495 550 TSP type-1 3.
FT DOMAIN 550 590 EGF-like 1.
FT DOMAIN 591 648 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 649 693 EGF-like 3.
FT DOMAIN 726 761 TSP type-3 1.
FT DOMAIN 762 784 TSP type-3 2.
FT DOMAIN 785 820 TSP type-3 3.
FT DOMAIN 821 843 TSP type-3 4.
FT DOMAIN 844 881 TSP type-3 5.
FT DOMAIN 882 917 TSP type-3 6.
FT DOMAIN 918 953 TSP type-3 7.
FT DOMAIN 954 1173 TSP C-terminal.
FT REGION 23 235 Heparin-binding (Potential).
FT MOTIF 929 931 Cell attachment site (Potential).
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 363 363 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 705 705 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 711 711 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1070 1070 N-linked (GlcNAc...) (Potential).
FT DISULFID 394 426 By similarity.
FT DISULFID 398 431 By similarity.
FT DISULFID 409 416 By similarity.
FT DISULFID 450 487 By similarity.
FT DISULFID 454 492 By similarity.
FT DISULFID 465 477 By similarity.
FT DISULFID 507 544 By similarity.
FT DISULFID 511 549 By similarity.
FT DISULFID 522 534 By similarity.

```

```
FT DISULFID 554 565 By similarity.
FT DISULFID 559 575 By similarity.
FT DISULFID 578 589 By similarity.
FT DISULFID 595 611 By similarity.
FT DISULFID 602 620 By similarity.
FT DISULFID 623 647 By similarity.
FT DISULFID 653 666 By similarity.
FT DISULFID 660 679 By similarity.
FT DISULFID 681 692 By similarity.
FT DISULFID 708 716 By similarity.
FT DISULFID 721 741 By similarity.
FT DISULFID 757 777 By similarity.
FT DISULFID 780 800 By similarity.
FT DISULFID 816 836 By similarity.
FT DISULFID 839 859 By similarity.
FT DISULFID 877 897 By similarity.
FT DISULFID 913 933 By similarity.
FT DISULFID 949 1170 By similarity.
SQ SEQUENCE 1173 AA; 130020 MW; A9F036D6516C0F24 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
DB 211 FQGVQLQNVRFV 221

RESULT 11
Q59E99 HUMAN
ID Q59E99 HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 21-FEB-2006, entry version 10.
DE Thrombospondin 1 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB209912; BAD93149.1; -; mRNA.
CC SMR; Q59E99; 886-939, 889-1225.
CC Ensembl; ENSG00000137801; Homo sapiens.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR006210; EGF_3.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR013032; EGF like reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSP_C.
```

```
DR InterPro: IPR001007; VWF_C.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSPREPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; WFC_1; 1.
DR PROSITE; PS0184; WFC_2; 1.
DR Cell adhesion; EGF-like domain.
KW NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16E57157B12 CRC64;

Query Match 100.0%; Score 55; DB 2; Length 1225;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
DB 263 FQGVQLQNVRFV 273

RESULT 12
Q5SPG5 BRARE
ID Q5SPG5 BRARE PRELIMINARY; PRT; 1090 AA.
AC Q5SPG5;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 21-FEB-2006, entry version 12.
DE Novel protein similar to vertebrate thrombospondin 1.\n (Fragment).
GN ORFNames=DKEY-11E23.1-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Barker D.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AL928866; CAI20599.1; -; Genomic_DNA.
CC SMR; Q5SPG5; 751-804, 754-1089.
CC Ensembl; ENSDARG0000010785; Danio rerio.
CC GO; GO:0005576; C:extracellular region; IEA.
CC GO; GO:0005509; F:calcium ion binding; IEA.
CC GO; GO:0008201; F:heparin binding; IEA.
CC GO; GO:0005515; F:protein binding; IEA.
CC GO; GO:0005198; F:structural molecule activity; IEA.
CC GO; GO:0007155; P:cell adhesion; IEA.
CC InterPro; IPR013320; ConA_like_subgrp.
CC InterPro; IPR002048; EF_hand_Ca_bd.
CC InterPro; IPR006210; EGF_3.
CC InterPro; IPR000742; EGF_3.
CC InterPro; IPR001881; EGF_Ca_bd.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR013032; EGF like reg.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
```

```
DR InterPro: IPR003367; tsp_3.  
DR InterPro: IPR008859; TSP_C.  
DR InterPro: IPR001007; VWF_C.  
DR Pfam: PF00008; EGF; 1.  
DR Pfam: PF00090; TSP_1; 2.  
DR Pfam: PF02412; TSP_3; 12.  
DR Pfam: PF05735; TSP_C; 1.  
DR Pfam: PF00093; VMC_1.  
DR PRINTS: PR01705; TSP1REPEAT.  
DR SMART: SM00181; EGF; 2.  
DR SMART: SM00209; TSP1; 2.  
DR SMART: SM00210; TSPN; 1.  
DR SMART: SM00214; VMC; 1.  
DR PROSITE: PS00018; EF_HAND_1; UNKNOWN_1.  
DR PROSITE: PS01186; EGF_2; 1.  
DR PROSITE: PS00266; EGF_3; 2.  
DR PROSITE: PS00092; TSP1; 2.  
DR PROSITE: PS01208; VWF_1; 1.  
DR PROSITE: PS0184; VWF_2; 1.  
KW Cell adhesion; EGF-like domain.  
FT NON_TER 1  
SQ SEQUENCE 1090 AA; 120978 MW; 5A9320504A22D836 CRC64;  
  
Query Match 90.9%; Score 50; DB 2; Length 1090;  
Best Local Similarity 90.9%; Pred. No. 1.1;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 FOGVLQNVRFV 11  
| | | | |  
Db 185 FMGVLQNVRFV 195  
  
RESULT 13  
ID Q4S758 TETNG PRELIMINARY; PRT; 1193 AA.  
AC Q4S758;  
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.  
DT 19-JUL-2005, sequence version 1.  
DE Chromosome 14 SCAF14723, whole genome shotgun sequence. (Fragment).  
OS ORFNames=GSTENG00022976001;  
GN Tetraodon nigroviridis (Tetraodon puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI_TaxID=99883;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15496914; DOI=10.1038/nature03025;  
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,  
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,  
RA Craud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,  
RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
the early vertebrate proto-karyotype."  
RL Nature 431:946-957(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RG Genoscope; Whitehead Institute Centre for Genome Research;  
RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
-----  
CC  
  
Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms  
Distributed under the Creative Commons Attribution-NoDerivs License  
-----  
CC EMBL; CAAB01014723; CAG03524.1; -; Genomic_DNA.  
CC SMR; Q4S758; 811-1148.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0008201; F:heparin binding; IEA.  
DR GO; GO:0005515; F:protein binding; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0007155; F:cell adhesion; IEA.  
DR InterPro: IPR006210; EGF.  
DR InterPro: IPR00742; EGF_3.  
DR InterPro: IPR001881; EGF_Ca bd.  
DR InterPro: IPR013032; EGF-like reg.  
DR InterPro: IPR003129; Laminin_G_TSP_N.  
DR InterPro: IPR00884; TSP1.  
DR InterPro: IPR008085; TSP_1.  
DR InterPro: IPR001007; VWF_C.  
DR Pfam: PF00008; EGF; 2.  
DR Pfam: PF00090; TSP_1; 3.  
DR Pfam: PF02412; TSP_3; 12.  
DR Pfam: PF05735; TSP_C; 1.  
DR Pfam: PF00093; VMC; 1.  
DR PRINTS: PR01705; TSP1REPEAT.  
DR SMART: SM00181; EGF; 2.  
DR SMART: SM00209; TSP1; 3.  
DR SMART: SM00210; TSPN; 1.  
DR SMART: SM00214; VMC; 1.  
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.  
DR PROSITE: PS00266; EGF_3; 2.  
DR PROSITE: PS00092; TSP1; 3.  
DR PROSITE: PS01208; VWF_1; 1.  
DR PROSITE: PS0184; VWF_2; 1.  
KW Cell adhesion.  
FT NON_TER 1  
SQ SEQUENCE 1193 AA; 133256 MW; 6E8781648FCEC7F2 CRC64;  
  
Query Match 90.9%; Score 50; DB 2; Length 1193;  
Best Local Similarity 90.9%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 FOGVLQNVRFV 11  
| | | | |  
Db 187 FMGVLQNVRFV 197  
  
RESULT 14  
Q5U903 PIG  
ID Q5U903 PIG PRELIMINARY; PRT; 249 AA.  
AC Q5U903;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 7.  
DE Thrombospondin 1 (Fragment).  
GN Name=Thbs1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;  
OC Sus.  
OX NCBI_TaxID=9823;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Zhang K., Maucio G., Hauet T.;  
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
-----  
CC  
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
-----  
CC EMBL; AY773342; AAV38110.1; -; mRNA.  
DR InterPro: IPR00884; TSP1.  
DR InterPro: IPR008085; TSP_1.  
DR InterPro: IPR001007; VWF_C.  
-----
```



```
DR Pfam: PF00090; TSP_1; 2.
DR Pfam: PF00093; VWC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
FT NON_TER 249 249
FT SEQUENCE 249 AA; 27560 MW; 465D664BE0329C32 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFV 11
Db 1 QGVQLQNVRFV 10

RESULT 15
Q4RLR5 TETNG PRELIMINARY; PRT; 1171 AA.
AC Q4RLR5_
DT 19-JUL-2005, integrated into UniprotKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Chromosome 10 SCAP15019, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSPENG00032374001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomerygia; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Briemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier J.-N., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAEE01015019; CAG10667.1; -; Genomic_DNA.
DR SNR; Q4RLR5; 834-887, 837-1171.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008201; F:heparin binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

```
DR InterPro: IPR006210; EGF.
DR InterPro: IPR000742; EGF_3.
DR InterPro: IPR001881; EGF_Ca_bd.
DR InterPro: IPR013032; EGF_like_reg.
DR InterPro: IPR003129; Laminin_G_TSP_N.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR008085; TSP_1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00090; TSP_1; 3.
DR Pfam: PF02412; TSP_3; 12.
DR Pfam: PF05735; TSP_C; 1.
DR Pfam: PF00093; VMC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Cell adhesion.
FT NON_TER 1171 1171
FT SEQUENCE 1171 AA; 129304 MW; 865F3749693F7FCE CRC64;

Query Match 89.1%; Score 49; DB 2; Length 1171;
Best Local Similarity 90.9%; Pred. No. 2;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FGVQLQNVRFV 11
Db 207 FGVQLQNVRFV 217

Search completed: June 5, 2006, 22:42:44
Job time : 100.853 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 22:43:07 ; Search time 21.9052 Seconds
(without alignments)
43.955 Million cell updates/sec

Title: US-10-030-735-30
Perfect score: 55
Sequence: 1 FQGVQLQVRV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Celerra_SID33/prodata/2/iaa/5 COMB.pep.*
- 2: /EMC Celerra_SID33/prodata/2/iaa/6 COMB.pep.*
- 3: /EMC Celerra_SID33/prodata/2/iaa/7 COMB.pep.*
- 4: /EMC Celerra_SID33/prodata/2/iaa/H COMB.pep.*
- 5: /EMC Celerra_SID33/prodata/2/iaa/PCUTUS COMB.pep.*
- 6: /EMC Celerra_SID33/prodata/2/iaa/RE COMB.pep.*
- 7: /EMC Celerra_SID33/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	825	2	US-09-949-002-482
2	55	100.0	831	2	US-09-939-853A-97
3	55	100.0	831	2	US-09-939-853A-98
4	55	100.0	1170	1	US-08-313-288B-20
5	55	100.0	1170	2	US-09-657-472-2
6	55	100.0	1170	2	US-09-949-002-350
7	39	70.9	731	2	US-09-270-767-42057
8	37	67.3	1045	2	US-09-949-016-11112
9	37	67.3	1172	1	US-08-313-288B-19
10	37	67.3	1172	2	US-09-949-016-6333
11	35	63.6	465	2	US-09-769-787-164
12	34	61.8	169	2	US-09-134-000C-3511
13	34	61.8	459	2	US-09-583-110-3429
14	34	61.8	715	1	US-08-484-393B-10
15	34	61.8	715	1	US-08-484-158B-10
16	34	61.8	715	1	US-08-484-596A-10
17	34	61.8	715	1	US-08-480-150A-10
18	34	61.8	715	2	US-08-458-731-10
19	34	61.8	715	2	US-08-149-223A-10
20	34	61.8	944	2	US-09-134-000C-5578
21	33	60.0	133	2	US-09-513-999C-7744
22	33	60.0	175	2	US-09-230-637-24
23	33	60.0	175	2	US-10-003-632C-5
24	33	60.0	390	2	US-09-328-352-4891
25	33	60.0	446	2	US-09-199-637A-267
26	33	60.0	484	2	US-09-248-796A-16180

27	33	60.0	538	2	US-09-134-000C-4331	Sequence 4331, Ap
28	33	60.0	743	2	US-09-708-392-1	Sequence 1, Appli
29	33	60.0	750	2	US-09-949-016-6004	Sequence 6004, Ap
30	33	60.0	750	2	US-09-949-016-6791	Sequence 6791, Ap
31	33	60.0	750	2	US-09-949-016-6792	Sequence 6792, Ap
32	33	60.0	750	2	US-09-949-016-6793	Sequence 6793, Ap
33	33	60.0	750	2	US-09-948-429C-63	Sequence 63, Appl
34	33	60.0	751	2	US-09-949-016-9533	Sequence 9533, Ap
35	33	60.0	751	2	US-09-949-016-9534	Sequence 9534, Ap
36	33	60.0	751	2	US-09-949-016-9535	Sequence 9535, Ap
37	33	60.0	751	2	US-09-949-016-9536	Sequence 9536, Ap
38	33	60.0	751	2	US-09-949-016-9537	Sequence 9537, Ap
39	33	60.0	751	2	US-09-949-016-9538	Sequence 9538, Ap
40	33	60.0	751	2	US-09-949-016-9539	Sequence 9539, Ap
41	33	60.0	751	2	US-09-949-016-9540	Sequence 9540, Ap
42	33	60.0	751	2	US-09-949-016-9541	Sequence 9541, Ap
43	33	60.0	751	2	US-09-949-016-9542	Sequence 9542, Ap
44	33	60.0	751	2	US-09-949-016-9543	Sequence 9543, Ap
45	33	60.0	751	2	US-09-949-016-9544	Sequence 9544, Ap

ALIGNMENTS

RESULT 1
US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-482

Query Match 100.0%; Score 55; DB 2; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQVRV 11
| | | | | | | | | |
DB 308 FQGVQLQVRV 318

RESULT 2
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232zel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCES: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20

```
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match      100.0%; Score 55; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQNVRV 11
      |||||
Db      208 FQGVQNVRV 218

RESULT 3
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Patent No. 6989232
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. 6989232el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/273,377
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-98

Query Match      100.0%; Score 55; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQNVRV 11
      |||||
Db      208 FQGVQNVRV 218

RESULT 4
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match      100.0%; Score 55; DB 1; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQNVRV 11
      |||||
Db      208 FQGVQNVRV 218

RESULT 5
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey Q.
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match      100.0%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQNVRV 11
      |||||
Db      208 FQGVQNVRV 218

RESULT 6
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
; US-09-949-002-350

Query Match      100.0%; Score 55; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQLQNVRFV 11
      |||||
Db      208 FQGVQLQNVRFV 218

RESULT 7
US-09-270-767-42057
; Sequence 42057, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42057
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-42057

Query Match      70.9%; Score 39; DB 2; Length 731;
Best Local Similarity 80.0%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 QGVQLQNVRFV 11
      |||||
Db      238 QGVQLQNVDFM 247

RESULT 8
US-09-949-016-11112
; Sequence 11112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11112
; LENGTH: 1045
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-11112

Query Match      67.3%; Score 37; DB 2; Length 1045;
Best Local Similarity 63.6%; Pred. No. 97;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 FQGVQLQNVRFV 11
      |||||
Db      281 FRGLLQNVHLV 291

RESULT 9
US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-288B-19

Query Match      67.3%; Score 37; DB 1; Length 1172;
Best Local Similarity 63.6%; Pred. No. 11e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 FQGVQLQNVRFV 11
      |||||
Db      202 FRGLLQNVHLV 212

RESULT 10
US-09-949-016-6333
; Sequence 6333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6333
; LENGTH: 1172
; TYPE: PR1
; ORGANISM: Human
US-09-949-016-6333

Query Match 67.3%; Score 37; DB 2; Length 1172;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
|:|:|:|:|
Db 202 FRGLQNVHLV 212

RESULT 11
US-09-769-787-164
; Sequence 164, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P211290
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 164
; LENGTH: 465
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-164

Query Match 63.6%; Score 35; DB 2; Length 465;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 10
|:|:|:|:|
Db 245 FNGTLENIRK 254

RESULT 12
US-09-134-000C-3511
; Sequence 3511, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3511
; LENGTH: 169
; TYPE: PR1
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (29)-(70)
; OTHER INFORMATION: Amino acids 29, 30, 59 & 70 are Xaa wherein Xaa = any amino acid
US-09-134-000C-3511

Query Match 61.8%; Score 34; DB 2; Length 169;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQNVR 9
|:|:|:|:|
Db 76 FQGVLRQIR 84

RESULT 13
US-09-583-110-3429
; Sequence 3429, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3429
; LENGTH: 459
; TYPE: PR1
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3429

Query Match 61.8%; Score 34; DB 2; Length 459;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
|:|:|:|:|
Db 132 FKATLENVQFV 142

RESULT 14
US-08-484-993B-10
; Sequence 10, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,993B
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 715 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-993B-10

Query Match 61.8%; Score 34; DB 1; Length 715;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVQLQVRP 10
| | | | |
Db 277 FPGKLSQVRP 286

RESULT 15
US-08-484-158B-10
Sequence 10, Application US/08484158B
Patent No. 5976545
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Pharmaceutical Compositions for
TITLE OF INVENTION: Immunocorrection
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 715 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-158B-10
Query Match 61.8%; Score 34; DB 1; Length 715;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 FQGVQLQVRP 10
| | | | |
Db 277 FPGKLSQVRP 286
Search completed: June 5, 2006, 22:48:55
Job time : 21.9052 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2006, 23:46:43 ; Search time 72.069 Seconds
(without alignments)
70.701 Million cell updates/sec

Title: US-10-030-735-30

Perfect score: 55

Sequence: 1 FQGVQLQVRV 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	12	4	US-10-474-213-28
2	55	100.0	240	4	US-10-419-462-40
3	55	100.0	240	5	US-10-782-968-40
4	55	100.0	432	5	US-10-741-600-1020
5	55	100.0	432	5	US-10-741-600-1022
6	55	100.0	459	6	US-11-043-806-462
7	55	100.0	466	3	US-09-925-301-1047
8	55	100.0	555	6	US-11-043-806-454
9	55	100.0	578	6	US-11-043-806-456
10	55	100.0	685	6	US-11-043-806-452
11	55	100.0	804	6	US-11-043-806-453
12	55	100.0	828	6	US-11-043-806-455
13	55	100.0	831	3	US-09-939-853A-97
14	55	100.0	831	3	US-09-939-853A-98
15	55	100.0	855	6	US-11-043-806-461
16	55	100.0	1000	6	US-11-043-806-457
17	55	100.0	1105	6	US-11-043-806-458
18	55	100.0	1150	4	US-10-296-733-1
19	55	100.0	1152	3	US-09-919-603-1
20	55	100.0	1169	5	US-10-317-821B-7
21	55	100.0	1170	4	US-10-020-141-12
22	55	100.0	1170	4	US-10-017-721-2
23	55	100.0	1170	4	US-10-021-660-114
24	55	100.0	1170	4	US-10-008-093-2
25	55	100.0	1170	4	US-10-295-027-1170
26	55	100.0	1170	4	US-10-211-462-38
27	55	100.0	1170	4	US-10-231-956A-482

28	55	100.0	1170	4	US-10-419-462-38	Sequence 38, Appl
29	55	100.0	1170	5	US-10-741-600-1018	Sequence 1018, Ap
30	55	100.0	1170	5	US-10-741-600-1019	Sequence 1019, Ap
31	55	100.0	1170	5	US-10-741-600-1021	Sequence 1021, Ap
32	55	100.0	1170	5	US-10-782-968-38	Sequence 38, Appl
33	55	100.0	1170	5	US-10-849-989-44	Sequence 44, Appl
34	55	100.0	1170	5	US-10-631-467-548	Sequence 548, App
35	55	100.0	1170	5	US-10-631-467-1376	Sequence 1376, Ap
36	55	100.0	1170	5	US-10-831-997-2	Sequence 2, Appli
37	55	100.0	1170	5	US-10-995-561-594	Sequence 594, App
38	55	100.0	1170	5	US-10-995-561-595	Sequence 595, App
39	55	100.0	1170	5	US-10-995-561-596	Sequence 596, App
40	55	100.0	1170	6	US-11-037-713-51	Sequence 51, Appl
41	55	100.0	1170	6	US-11-046-644-28	Sequence 28, Appl
42	55	100.0	1170	6	US-11-046-456-28	Sequence 28, Appl
43	45	81.8	15	4	US-10-285-394-153	Sequence 153, App
44	39	70.9	226	5	US-10-467-657-2428	Sequence 2428, Ap
45	39	70.9	721	6	US-11-097-143-13287	Sequence 13287, A

ALIGNMENTS

RESULT 1

US-10-474-213-28

; Sequence 28, Application US/10474213

; Publication No. US20040214248A1

; GENERAL INFORMATION:

; APPLICANT: Roberts, David D

; APPLICANT: Kruttsch, Henry C

; TITLE OF INVENTION: USE OF SEMENOGELIN IN THE DIAGNOSIS, PROGNOSIS AND TREATMENT OF

; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 224329

; CURRENT APPLICATION NUMBER: US/10/474,213

; CURRENT FILING DATE: 2003-10-06

; PRIOR APPLICATION NUMBER: PCT/US02/10535

; PRIOR FILING DATE: 2002-04-03

; PRIOR APPLICATION NUMBER: 60/281,994

; PRIOR FILING DATE: 2001-04-06

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 28

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Synthetic peptide

US-10-474-213-28

Query Match 100.0%; Score 55; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQVRV 11

Db 1 FQGVQLQVRV 11

RESULT 2

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Williams, Kevin J.

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 40

; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-419-462-40

Query Match 100.0%; Score 55; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.025; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
|||
Db 190 FQGVQNVRV 200

RESULT 3

US-10-782-968-40
; Sequence 40, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; Cancellation of Patent: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain
US-10-782-968-40

Query Match 100.0%; Score 55; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.025; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
|||
Db 190 FQGVQNVRV 200

RESULT 4

US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 100.0%; Score 55; DB 5; Length 432;

Best Local Similarity 100.0%; Pred. No. 0.048; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FQGVQNVRV 11
|||
Db 208 FQGVQNVRV 218

RESULT 5

US-10-741-600-1022
; Sequence 1022, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1022
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(432)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 100.0%; Score 55; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.048; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
|||
Db 208 FQGVQNVRV 218

RESULT 6

US-11-043-806-462
; Sequence 462, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Method
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 462
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-462

Query Match 100.0%; Score 55; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.051; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRV 11
|||
Db 208 FQGVQNVRV 218

RESULT 7

US-09-925-301-1047
; Sequence 1047, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1047
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 100.0%; Score 55; DB 3; Length 466;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
| | | | | | | | | |
Db 261 FQGVQNVRFV 271

RESULT 8
US-11-043-806-454
; Sequence 454, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 454
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-454

Query Match 100.0%; Score 55; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 0.063;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
| | | | | | | | | |
Db 208 FQGVQNVRFV 218

RESULT 9
US-11-043-806-456
; Sequence 456, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 456
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-456

Query Match 100.0%; Score 55; DB 6; Length 578;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
| | | | | | | | | |
Db 208 FQGVQNVRFV 218

RESULT 10
US-11-043-806-452
; Sequence 452, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 452
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-452

Query Match 100.0%; Score 55; DB 6; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
| | | | | | | | | |
Db 208 FQGVQNVRFV 218

RESULT 11
US-11-043-806-453
; Sequence 453, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 453
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-453

Query Match 100.0%; Score 55; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
| | | | | | | | | |
Db 208 FQGVQNVRFV 218

RESULT 12
US-11-043-806-455
; Sequence 455, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 455

```
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-455

Query Match      100.0%; Score 55; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
Db 208 FQGVQLQNVRFV 218

RESULT 13
US-09-939-853A-97
; Sequence 97, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match      100.0%; Score 55; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
Db 208 FQGVQLQNVRFV 218

RESULT 14
US-09-939-853A-98
; Sequence 98, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-09-939-853A-98

Query Match      100.0%; Score 55; DB 3; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
Db 208 FQGVQLQNVRFV 218

RESULT 15
US-11-043-806-461
; Sequence 461, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043,806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 461
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-461

Query Match      100.0%; Score 55; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |||||
Db 208 FQGVQLQNVRFV 218

Search completed: June 6, 2006, 00:00:11
Job time : 72.069 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2006, 00:00:38 ; Search time 3.41379 Seconds
(without alignments)
37.266 Million cell updates/sec

Title: US-10-030-735-30

Perfect score: 55

Sequence: 1 QGVQLQNVRFV 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /EMC_Celerra_SID33/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
- 2: /EMC_Celerra_SID33/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
- 3: /EMC_Celerra_SID33/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
- 4: /EMC_Celerra_SID33/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
- 5: /EMC_Celerra_SID33/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
- 6: /EMC_Celerra_SID33/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
- 7: /EMC_Celerra_SID33/ptodata/1/pubpaa/US11_NEW_PUB.pap.*
- 8: /EMC_Celerra_SID33/ptodata/1/pubpaa/US60_NEW_PUB.pap.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	60.0	298	US-10-953-349-9347	Sequence 9347, A
2	33	60.0	331	US-10-953-349-34674	Sequence 34674, A
3	33	60.0	358	US-10-953-349-34673	Sequence 34673, A
4	33	60.0	382	US-10-953-349-34672	Sequence 34672, A
5	33	60.0	750	US-10-511-937-2413	Sequence 2413, A
6	32	58.2	843	US-10-953-349-16527	Sequence 16527, A
7	32	58.2	962	US-10-953-349-16526	Sequence 16526, A
8	32	58.2	1014	US-10-953-349-16525	Sequence 16525, A
9	31	56.4	608	US-11-203-828-5	Sequence 5, Appli
10	31	56.4	855	US-11-247-437-2	Sequence 3713, A
11	30	54.5	132	US-10-953-349-39312	Sequence 39138, A
12	30	54.5	178	US-11-293-697-3713	Sequence 39138, A
13	30	54.5	227	US-10-953-349-39318	Sequence 39138, A
14	30	54.5	251	US-10-953-349-3586	Sequence 3586, A
15	30	54.5	265	US-10-953-349-39137	Sequence 39137, A
16	30	54.5	276	US-10-953-349-39136	Sequence 39136, A
17	30	54.5	277	US-10-953-349-3585	Sequence 3585, A
18	30	54.5	329	US-10-953-349-24690	Sequence 24690, A
19	30	54.5	332	US-10-953-349-24689	Sequence 24689, A
20	30	54.5	404	US-10-953-349-24688	Sequence 24688, A
21	30	54.5	407	US-10-953-349-26200	Sequence 26200, A
22	30	54.5	724	US-11-293-697-3263	Sequence 3263, A
23	30	54.5	1333	US-10-511-937-2992	Sequence 2992, A
24	29	52.7	139	US-11-293-697-3323	Sequence 3323, A
25	29	52.7	230	US-10-953-349-28990	Sequence 28990, A

ALIGNMENTS

RESULT 1

US-10-953-349-9347
; Sequence 9347, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9347
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9347

Query Match 60.0%; Score 33; DB 6; Length 298;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFV 11
DB 87 EGLIQGVRFI 96

RESULT 2

US-10-953-349-34674
; Sequence 34674, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34674
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34674

Query Match 60.0%; Score 33; DB 6; Length 331;

```
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFV 11
   |||| |::|
Db 225 QGVLFNIQYV 234
   |||| |::|

RESULT 3
US-10-953-349-34673
; Sequence 34673, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34673
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34673

Query Match 60.0%; Score 33; DB 6; Length 358;
Best Local Similarity 60.0%; Pred. No. 18;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFV 11
   |||| |::|
Db 252 QGVLFNIQYV 261
   |||| |::|

RESULT 4
US-10-953-349-34672
; Sequence 34672, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34672
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34672

Query Match 60.0%; Score 33; DB 6; Length 382;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFV 11
   |||| |::|
Db 276 QGVLFNIQYV 285
   |||| |::|

RESULT 5
US-10-511-937-2413
; Sequence 2413, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
```

```
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2413
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2413

Query Match 60.0%; Score 33; DB 6; Length 750;
Best Local Similarity 40.0%; Pred. No. 40;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRF 10
   |::|::|
Db 508 FENIIQLNKF 517
   |::|::|

RESULT 6
US-10-953-349-16527
; Sequence 16527, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16527
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16527

Query Match 58.2%; Score 32; DB 6; Length 843;
Best Local Similarity 54.5%; Pred. No. 72;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQLQNVRFV 11
   |::|::|
Db 780 FKGILDNVYFV 790
   |::|::|

RESULT 7
US-10-953-349-16526
; Sequence 16526, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
```

; SEQ ID NO 16526
; LENGTH: 962
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16526

Query Match 58.2%; Score 32; DB 6; Length 962;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
|:|:|:|:|
Db 899 FKGLDNNYVFV 909

RESULT 8

US-10-953-349-16525
; Sequence 16525, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16525
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-16525

Query Match 58.2%; Score 32; DB 6; Length 1014;
Best Local Similarity 54.5%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
|:|:|:|:|
Db 951 FKGLDNNYVFV 961

RESULT 9

US-11-203-828-5
; Sequence 5, Application US/11203828
; Publication No. US20060110390A1
; GENERAL INFORMATION:
; APPLICANT: LEINWARD, LESLIE
; APPLICANT: SUCHAROV, CARMEN
; TITLE OF INVENTION: INHIBITION OF KU AS A TREATMENT FOR CARDIOVASCULAR DISEASES
; FILE REFERENCE: MYOG:58US
; CURRENT APPLICATION NUMBER: US/11/203,828
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/604,435
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-203-828-5

Query Match 56.4%; Score 31; DB 7; Length 608;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQNVRF 10
: : : : :
Db 461 KAIVQNVRF 469

RESULT 10

US-11-247-437-2
; Sequence 2, Application US/11247437
; Publication No. US20060110753A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Roberts, Bruce
; TITLE OF INVENTION: METHODS TO DIAGNOSE AND TREAT LUNG CANCER
; FILE REFERENCE: 5257C
; CURRENT APPLICATION NUMBER: US/11/247,437
; CURRENT FILING DATE: 2005-10-11
; PRIOR APPLICATION NUMBER: PCT/US2004/011193
; PRIOR FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US 60/462,028
; PRIOR FILING DATE: 2003-04-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-247-437-2

Query Match 56.4%; Score 31; DB 7; Length 855;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVQNVRFV 11
|:|:|:|:|
Db 445 RGVYENVKIV 454

RESULT 11

US-10-953-349-39312
; Sequence 39312, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39312
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-39312

Query Match 54.5%; Score 30; DB 6; Length 132;
Best Local Similarity 45.5%; Pred. No. 23;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVQNVRFV 11
|:|:|:|:|
Db 66 FEGILNNNVFI 76

RESULT 12

US-11-293-697-3713
; Sequence 3713, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28

```
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3713
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3713

Query Match          54.5%; Score 30; DB 7; Length 178;
Best Local Similarity 45.5%; Pred. No. 32;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 FQGVQLQVRV 11
DB      4 FQAVVNLNFI 14

RESULT 13
US-10-953-349-39138
; Sequence 39138, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39138
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-39138

Query Match          54.5%; Score 30; DB 6; Length 227;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 LQNVRFV 11
DB      183 LRNVRFV 189

RESULT 14
US-10-953-349-3586
; Sequence 3586, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3586
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3586

Query Match          54.5%; Score 30; DB 6; Length 251;
Best Local Similarity 45.5%; Pred. No. 47;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 FQGVQLQVRV 11
DB      80 FQGHENLPPL 90
```

```
RESULT 15
US-10-953-349-39137
; Sequence 39137, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39137
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-39137

Query Match          54.5%; Score 30; DB 6; Length 265;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 LQNVRFV 11
DB      221 LRNVRFV 227

Search completed: June 6, 2006, 00:12:56
Job time : 3.51379 secs
```